EXPLAINING RARE MENDELIAN PHENOTYPES: EXOME SEQUENCING AND FUNCTIONAL ANALYSIS OF SPONDYLOMETAPHYSEAL DYSPLASIA WITH CONE-ROD DYSTROPHY

by
Julie Jurgens

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
December, 2017

Abstract

Mendelian disorders are a set of heritable phenotypes that come at a significant cost for patients, their families, and the healthcare system. Here we describe and implement strategies to determine the genetic basis of unexplained Mendelian disorders through next-generation sequencing. Using whole-exome sequencing, we found that a novel heterozygous variant in *COL2A1* underlies Stanescu dysplasia, and that biallelic variants in *PCYT1A* cause spondylometaphyseal dysplasia with cone-rod dystrophy (SMD-CRD). In addition to detecting causal variants for phenotypes of interest, next-generation sequencing has the potential to reveal incidental findings in genes unrelated to a patient's primary phenotype. We find that incidental findings arise frequently in clinically actionable genes, yet our ability to interpret them is extremely limited, highlighting the need for improved and standardized variant classification methods. Finally, we generated and characterized various cellular models for the characterization of SMD-CRD alleles in *PCYT1A*. Our studies suggest that full or partial loss of *PCYT1A* expression has variable and cell-type specific consequences, resulting in altered lipid metabolism and lipid droplet formation, accelerated chondrocyte differentiation, and varying degrees of *PCYT1A* loss-of-function. We conclude that concerted efforts are required to fully elucidate the genetic and mechanistic bases of Mendelian phenotypes, and that sustained efforts are required for the accurate interpretation of human genetic variation detected by next-generation sequencing studies.

Thesis Advisor: Dr. David Valle

Thesis Readers: Drs. David Valle and Donald Zack

Preface

I owe the Human Genetics Graduate Program tremendous thanks for investing in me and in my graduate education. In many ways, I feel as though I've learned more in these past five years than in my entire pre-graduate school life, and witnessing the fascinating advances by colleagues and friends has been a constant impetus for me to push forward with my own research. The Human Genetics Program and all its affiliated faculty and staff, particularly the members of the executive committee (Drs. David Valle, Kirby Smith, Andy McCallion, Susan Michaelis, and Roger Reeves, and Ms. Sandy Muscelli) have been constant sources of support throughout my graduate career and have made completion of my degree possible. Thanks to this program, I am certain that genetics is my calling and couldn't be more excited to shape a career in this area.

Many kind and selfless individuals have helped me through graduate school, and I would like to thank them all for their commitment to my development as a scientist and as a person. I would especially like to thank my advisor, Dr. David Valle, for his constant encouragement and support. Dr. Valle has taught me how to think like a geneticist; how to maintain curiosity, passion, and persistence despite experimental misfortunes; and how to contextualize my work in terms of what is known and what still remains to be discovered. I'm grateful to have had an advisor who is not just a supervisor, but a mentor, and who takes every opportunity to teach me, encourage my development, and to grant me the intellectual guidance as well as freedom I need to develop as an independent research investigator. I could not have been luckier in choosing Dr. Valle as my advisor, and hope I will one day be able to fill even half his footsteps.

In addition to having a wonderful advisor, I've had a fantastic thesis committee that has bolstered my scientific prowess and given me tremendous opportunities for intellectual and professional development. First, Dr. Daniel Raben has served as chair of my thesis committee and has been a fount of knowledge for all aspects of lipid biology. He was kind enough to invite me into his lab space so I could do my lipid work, and even went so far as to teach me bench techniques himself when other lab members were unavailable. Though I still consider myself more a geneticist than a lipid biologist, I am grateful for the wholehearted welcome to Lipid Land. I've appreciated the opportunity to more deeply examine the functional consequence of human variation thanks to his guidance, as well as his genuine kindness and concern for my development as a student throughout graduate school. Becky Tu-Sekine in the Raben lab has also been a wonderful resource, and I am grateful for all her guidance in experimental lipid techniques. Secondly, Dr. Donald Zack has been an invaluable member of my committee, providing crucial insights into cell biology and the visual sciences. Dr. Zack is not only a rigorous and innovative scientist, but also has always gone out of his way to support my development as a researcher and even offered to serve as one of my thesis readers. Dr. Zack gave me the opportunity to deepen my understanding of visual systems through the Visual Neuroscience Training Program (VNTP), which will be the foundation of my work as a postdoctoral fellow. I'm extremely thankful to Dr. Zack for all his kindness and support over the years. And finally, Dr. Nara Sobreira has been a mentor and role model throughout graduate school, from training me to analyze exome sequencing data during my rotation in the lab to giving me the opportunity to collaborate on a variety of next-generation sequencing analysis projects through the Baylor-Hopkins Center for

Mendelian Genomics. I've greatly appreciated her guidance, encouragement, and training over the years, and in many ways she has emboldened me to become an independent scientist.

I would also like to thank Ms. Sandy Muscelli for her tireless efforts in guiding me through the program. Sandy has always provided honest, thoughtful advice, a listening ear, and incredible kindness. She's the one who forwarded me the advertisement for my future postdoc position, and has given me many additional opportunities in my time here. I'll always be grateful for all her efforts to ensure the success of students in the program, and can't thank her enough for all she's done for me.

During grad school, I've had the privilege of working with a number of collaborators in both the clinical and research realms, and would like to acknowledge their contributions to this work and to my intellectual and professional growth. Particular thanks go to Drs. Julie Hoover-Fong, Rosemary Cornell, Norman Haughey, Suming Chen, Raha Dastgheyb, Saja Khuder, Jill Fahrner, Hilary Vernon, and Courtney Woods for my functional studies of *PCYT1A* variants; Drs. Richard Pauli and Tae-Joon Cho and their colleagues, whom I had the pleasure to work with on my *COL2A1* project; Drs. Jeremy Nathans, Martha Brucato, Dimitri Avramopoulos, and Christine Ladd-Acosta for the study of strabismus; Drs. Joann Bodurtha, Hal Dietz, Joel Hirschorn, Brendan Lee, Mike Guo, and Mahim Jain as well as Ms. Weiyi Mu and Ms. Gretchen Oswald for my Ehlers-Danlos project; all the collaborators for my incidental findings project; Dr. Garry Cutting and Ms. Karen Raraigh for my atypical cystic fibrosis project; Dr. Kathy Wilson, Ms. Alyssa Florwick and Mr. Tejas Dharmaj for my *LMNA* project; Drs. Salmo Raskin, Anna Schossig, and Ines Kapferer-Seebacher for my *SLC13A5* project; Drs. Julie

Hoover-Fong and Ada Hamosh for their collaboration on multiple projects and for keeping the clinical aspects of our research moving forward; Drs. Guilerme Yamamoto, Felicity Collins, John Christodoulou, Mike Bober, and Lakshmi Mehta for clinical follow-up for SMD-CRD patients and for providing patient skin biopsies; and Drs. Mary Armanios and Jing You for the functional analysis of *TELO2* variants uncovered in one of my exome sequencing analyses. And finally, I'd like to thank all the other individual collaborators I've been able to work with on individual families' exome analyses through the Baylor-Hopkins Center for Mendelian Genomics.

I also owe a tremendous thanks to the team at the Center for Inherited Disease Research, particularly Drs. Kim Doheny, Hua Ling, Dane Witmer, Kurt Hetrick, Elizabeth Pugh, and Peng Zhang. They have been not only extremely helpful in conducting data analyses for our sequencing studies, but also in helping me develop skills to do these analyses independently and in referring me to new resources that may be of use for my work. I am extremely grateful for all their help and kindness over the years.

The Valle lab has been my family way from home for the last few years, and I've been incredibly fortunate to have such wonderful colleagues and friends there throughout graduate school. Cassandra Obie and Gary Steel have been a wealth of information for all my cell culture and molecular biology experiments, as well as steadfast sources of support, advice, and laughter. My fellow graduate students in the lab, Sarah Robbins and Jing You, have been tremendously compassionate, supportive, and fun, and I've learned a great deal from both of them. Furthermore, members of the Avramopoulos, Vernon, and Sobreira labs have been like my extended lab family and have helped me in many ways; particularly I would like to thank Nicole Eckart, Xuan Pham, Ruihua Wang, Kyra Feuer,

Michele Migliavacca, Renan Martin, Elizabeth Wohler, Mike Muriello, and Arianna Franca.

I owe tremendous thanks as well to the Visual Neuroscience Training Program for funding my fellowship and assisting me with learning about the visual sciences through coursework and seminars during my time here. I would also like to thank the Burroughs-Wellcome Fund/ Maryland Genetics, Epidemiology, and Medicine Program for granting Martha Brucato and me the 2016 Wolfe Street Competition Award in order to study the genetic basis of strabismus. I've learned a great deal and have been able to pursue my work more thoroughly thanks to these opportunities, and am very grateful for them.

Finally, I would like to thank my family, friends, and former colleagues for their unwavering support throughout this process. I've been surrounded by role models and positive influences in every stage of my life, and couldn't be more grateful. I owe particular thanks to my parents, Carol Cunningham and Carl Jurgens, for their constant encouragement, love, and support.

Table of Contents

List of Tables

Chapter 2 Tables:

Chapter 3 Tables:

Chapter 4 Tables:

List of Figures

Chapter 2 Figures:

Chapter 3 Figures:

Chapter 4 Figures:

Chapter 5 Figures:

Chapter 1: Introduction

The utility of studying rare Mendelian disorders

Mendelian disorders are a collection of heritable phenotypes transmitted from parents to their children according to Mendelian modes of inheritance. Approximately 8% of all viable human births have a Mendelian condition that manifests by adulthood, and Mendelian disorders affect a total of 25 million Americans (Chong et al., 2015). Many of these disorders are extremely debilitating and are of high burden for patients, their families, and the healthcare system (Angelis, Tordrup, & Kanavos, 2015). Mendelian disorders range from severe, multi-systemic phenotypes such as Hutchinson-Gilford progeria syndrome to milder, more targeted phenotypes such as isolated strabismus. Mendelian disorders vary in their population frequencies, and in some instances the same disorder can vary in frequency depending on the population being studied, as pathogenic alleles may be enriched in specific populations. An example of this is Tay-Sachs disease, for which 1:300 individuals in the general population are heterozygous carriers of pathogenic alleles, in contrast to the 1:30 heterozygous carrier rate in the Ashkenazi Jewish population (MIM 272800; www.omim.org).

Because many monogenic Mendelian disorders affect only small numbers of individuals, the utility of studying them comes into question, especially when research resources are limited. As a form of justification, we can think of Mendelian phenotypes as naturally occurring model systems. They enable us to see how an isolated genetic defect can impact larger biological processes, and in this way dissect systems which would otherwise be far too complex to understand all at once. Furthermore, studying monogenic extremes sometimes provide insights which are more broadly applicable to

the population at large. An example of this is Brown and Goldstein's discovery of biallelic variants in the gene encoding the low-density lipoprotein receptor in individuals with familial hypercholesterolemia (Goldstein & Brown, 2015). By studying this severe pediatric disorder, Brown and Goldstein were not only able to dissect more generalizable mechanisms for LDL uptake and cholesterol metabolism, but also were able to implicate heterozygous LDL receptor variants in less severe cardiovascular phenotypes affecting 1:500 individuals. Finally, the basic research that stemmed from uncovering this genetic defect led to the discovery and implementation of statin drugs, which have been used to reduce LDL levels and prevent cardiovascular disease in the population at large. Thus, understanding extreme cases that occur naturally in the population can ultimately lead to broader clinical applications and help us to understand genetic variation as a whole.

Application of whole-exome and whole-genome sequencing to explain rare Mendelian disorders

Next-generation sequencing has ushered in a brave new era for diagnosis and treatment of Mendelian disorders. Previously, disease gene discovery relied on hypothesis-driven testing for disease-associated genes and pathogenic variants, and as such was biased toward existing knowledge of variation in humans and other model systems. Exome and genome sequencing enable unbiased evaluation of genetic variation in order to identify candidate causal variants, and thus enable the discovery of completely novel mechanisms for disease. Less than a decade ago, Ng et al. demonstrated that exome sequencing can be applied to uncover the genetic basis of a Mendelian disorder (S. B. Ng et al., 2010) and opened the floodgates for applying exome sequencing on a broader scale.

In 2012, the NIH, NHGRI, and NHLBI collectively initiated the Centers for Mendelian Genomics (CMGs), a collection of centers across the US dedicated to finding the genetic bases of unexplained Mendelian disorders through next-generation sequencing. One of the CMGs is the Baylor-Hopkins Center for Mendelian Genomics, led by investigators at Johns Hopkins University and Baylor University. As an initial step toward identifying all genetic variants and genes contributing to Mendelian phenotypes, the CMGs began by applying exome sequencing to understand the genetic basis of a variety of Mendelian disorders. Although the exome only comprises ~1-2% of the genome, variants in protein-coding sequences are responsible for the majority of explained Mendelian disorders to-date (Chong et al., 2015). Furthermore, exome sequencing is currently less expensive than genome sequencing, so a larger number of phenotypes can be investigated by applying this approach as an initial step. As of December 16, 2017, http://mendelian.org/mendelian-traits-numbers stated that 63% of currently known Mendelian phenotypes have a known molecular basis, and 18% of human genes have been shown to harbor variants that cause a Mendelian phenotype. Thus, we are making great headway in understanding Mendelian phenotypes and their genetic bases, but much of the exome still remains unexplained.

Hurdles for gene and variant discovery and functional characterization

Though next-generation sequencing together with analytic approaches that incorporate genetic and genomic information have solved many Mendelian disorders, several factors have prevented complete success of these studies. First, these disorders typically are rare, which prevents obtaining sufficient statistical evidence for genetic causality, and pedigrees large enough for convincing segregation analysis are difficult to

find. In some cases, unaffected or more mildly affected family members may also be mosaic for variants of interest, which confounds analyses. Second, our ability to phenotype is limited by the systems subjected to clinical examination, the expertise and thoroughness of examination, and the age of the patients being studied. In some cases, phenotypes are missed simply because clinicians don't yet know to look for them, or because they have not yet manifested in a clinically recognizable way. Both these factors confound our ability to cluster families for analytical purposes, or to perform proper segregation analyses within families for late-onset conditions.

Moreover, there is often phenotypic overlap among genetically disparate disorders (for instance, joint hypermobility in multiple forms of Ehlers-Danlos as well as other disorders). Another issue is locus heterogeneity (variants in different genes causing similar phenotypes), since the same candidate gene may be mutated in only a single family. Furthermore, allelic heterogeneity (different variants in the same gene observed in independent individuals with the same phenotype) complicates variant interpretation, as many variants fall in disease-associated genes but are not pathogenic in themselves. We will explore this concept further in Chapter 3.

Nearly a quarter of all phenotype-associated genes have been described in connection with more than one Mendelian phenotype, which further confounds our ability to interpret variants (December 16, 2017, http://mendelian.org/mendelian-traits-numbers). I will provide an example of this in Chapter 5 through the case of biallelic variants in *PCYT1A* causing three distinct or partially overlapping phenotypes. A separate problem is that some disorders appear to follow classic Mendelian patterns of inheritance, but are not monogenic; this can also contribute to reduced penetrance and

variable expressivity, which pose additional challenges for segregation analyses. Furthermore, genes may require environmental interactions in order for protection against disease or to spur disease manifestation (e.g. fava bean consumption leading to hemolytic anemia in association with G6PD deficiency; MIM 305900).

Another hurdle is obtaining strong evidence for variant causality, particularly given a large list of candidate variants. For instance, in conducting exome sequencing of an affected parent and child with an autosomal dominant disorder, we may find ~100 rare (minor allele frequency < 0.01), functional (coding or splice site) variants shared between both affected individuals. I will discuss such a case in Chapter 2. Some genes such as *TTN* also have large coding sequences, and as such are larger targets for mutation. This makes it extremely difficult to interpret variation that arises in these genes. In some instances, certain variants may also be in linkage disequilibrium with truly pathogenic variants, which can lead to improper flagging of variants as causal. Finally, strong candidate variants may be identified but lack additional supporting functional evidence to implicate them in disease pathogenesis. This issue can be compounded by inaccessibility of relevant tissue types for certain disorders. Finally, although exome sequencing enriches for protein-coding regions, it excludes putatively functional noncoding elements such as promoters or enhancers, as well as much of the 3' and 5' untranslated regions and introns. Our ability to detect copy number variants as well as structural variants from exomes is also limited. In combination, all these factors have barred us from having a more complete understanding of Mendelian disorders and their genetic bases. However, they have also provided lessons for conducting future analyses more effectively and innovatively.

Finally, once candidate variants have been identified, detailed functional studies must be conducted in order to trace the path from a genetic defect to its corresponding disease mechanism. Although this process may be straightforward, biological systems are often more complex than we anticipate. The consequences of genetic variants may be tissue- or cell-type-specific, and disease mechanisms may not be overtly connected to the known functions of candidate genes, as we will see in Chapter 5. Though riddled with challenges, increasingly sophisticated methods for functional studies are necessary for truly understanding Mendelian disorders and treating them to improve patient outcomes. Next-generation sequencing studies have invigorated a whole new cycle of functional interrogation and biological understanding, which will only continue to grow as we move forward.

Chapter 2: Novel *COL2A1* variant (c.619G>A, p.Gly207Arg) manifesting as a phenotype similar to progressive pseudorheumatoid dysplasia and spondyloepiphyseal dysplasia, Stanescu type

Chapter 2 is reprinted from the final accepted version of the journal article:

Jurgens J, Sobreira N, Modaff P, Reiser CA, Seo SH, Seong MW, Park SS, Kim OH, Cho TJ, Pauli RM. (2015). Novel *COL2A1* variant (c.619G>A, p.Gly207Arg) manifesting as a phenotype similar to progressive pseudorheumatoid dysplasia and spondyloepiphyseal dysplasia, Stanescu type. *Hum Mutat.* 36(10): 1004-8.

Julie Jurgens's Contributions: Exome sequence analysis and identification of *COL2A1* as primary candidate gene for this family's condition; Sanger validation of variant and segregation analysis; identification of second family with the same *COL2A1* variant and skeletal dysplasia phenotype at an ASHG annual meeting; coordination of efforts between clinicians and researchers in U.S. and Korea for phenotypic comparison of probands, correspondence, and ongoing collaboration; conceptualization and drafting of manuscript and figures; coordination of revisions from other co-authors; coordination of submission process; drafting of response to reviewer comments and integration of additional responses from other co-authors; manuscript finalization with journal for publication; generation of graphical abstract as requested by journal

Introduction

Progressive pseudorheumatoid dysplasia (PPRD; MIM 208230), is a painful, disabling autosomal recessive skeletal dysplasia accompanied by pseudorheumatoid arthritis that manifests in childhood and inexorably worsens. Features include platyspondyly, epimetaphyseal expansion, and progressive stiffness and swelling of the joints (Garcia Segarra et al., 2012). Hurvitz et al. (1999) implicated variants in *WISP3* (Wnt-1-inducible secreted protein 3; MIM 603400) in several families with PPRD; however, no *WISP3*

variants were found in a simplex case in this cohort (Family 6), suggesting locus heterogeneity (Hurvitz et al., 1999).

Spondyloepiphyseal dysplasia (SED), Stanescu type, is distinguished by painful, progressive joint space narrowing, joint contractures, platyspondyly, and epimetaphyseal dysplasia with pronounced femoral coxa valga and normal stature. This disorder was first described by Stanescu et al. (1984) in a male proband born of unaffected parents, making the mode of inheritance indeterminate. Subsequently, Nishimura et al. (1998) described two families with multiple affected generations and autosomal dominant transmission (Nishimura et al., 1998). As noted by Nishimura et al. (1998), SED, Stanescu type is likely underdiagnosed, perhaps due to shared features with other dysplasias. Neither the genetic basis of this disorder nor is its relationship to PPRD is clear.

Czech dysplasia (MIM 609162) is caused by a heterozygous c.823C>T, p.Arg275Cys variant in the type II alpha-1 collagen gene *COL2A1* (MIM 120140). While also characterized by early-onset, progressive arthritis and skeletal dysplasia, it is clinically distinguishable from PPRD by shortened metatarsals in all described individuals (Marik, Marikova, Zemkova, Kuklik, & Kozlowski, 2004).

Here we describe three patients from two families harboring a novel heterozygous variant in *COL2A1* (RefSeq NM_001844.4: c.619G>A, p.Gly207Arg), with similar phenotypes.

Clinical Description

The proband of Family 1 (Table 1) was previously reported as the simplex case of family 6 by Hurvitz et al. (1999). She is the product of unaffected, non-consanguineous parents with negative family histories for relevant features. As a child, she was noted

never to hop or run. A waddling gait was initially assessed at 4 years of age, prompting concerns of limb weakness, but extensive evaluation demonstrated no neuromuscular abnormality. Joint pain began around 5 years of age. More severe joint pain and early joint immobility arose in late childhood and were initially most severe in her neck and spine.

Although radiographs at 9 years of age showed mild platyspondyly, anterior vertebral wedging and mild, generalized epiphyseal abnormalities, she carried a diagnosis of juvenile rheumatoid arthritis throughout childhood. She developed kyphoscoliosis in adolescence and had extension osteotomies of both femora at age 13 years due to worsening hip flexion contractures.

She was first evaluated through the Midwest Regional Bone Dysplasia Clinics at age 22 years (Fig. 1a-c). Between childhood and her first assessment, joint pain progressively worsened, with greatest severity in her elbows, hips, knees and hands. Progressive joint stiffness and joint prominence developed, particularly in her hands. She had an adult height of ~150 cm, generalized truncal stiffness, severe neck stiffness, normal craniofacial features, and limited mouth opening. Shoulder abduction and elevation as well as elbow extension, pronation and supination were limited. Legs showed hip flexion contractures, only 15° of rotation, no abduction or adduction, minimal active hip flexion, mild knee flexion contractures, and ankle stiffness. Her stance was a "Z" posture (Fig. 1a-b) with absent heel strike, knee flexion, and hip flexion. Spine radiographs showed platyspondyly and anterior wedging (Fig. 1c).

Over the ensuing ten years, her stiffness and pain worsened, requiring bilateral hip replacements at age 25 years and knee replacements at age 26 years. She developed

progressive thoracolumbar kyphosis and presented acutely at age 35 years with progressive numbness to the level of the umbilicus and leg weakness. MRI showed severe spinal stenosis and multilevel myelomalacia. She had laminectomy (T8-L1 and L4-L5) and spinal fusion, which improved her symptoms. Today, she is 38 years old and remains ambulatory.

She was initially diagnosed with PPRD and included in the cohort of Hurvitz et al. (1999), but had no pathogenic *WISP3* variants, which was confirmed by subsequent testing (Supp. Materials and Methods). A diagnosis of Czech dysplasia was also considered but rejected due to the absence of structural abnormalities of the toes or metatarsals.

Concerns arose in the proband's daughter (Table 1) during childhood. She was born at 36 weeks gestation with a weight of 2840 g and length of 48 cm. When first assessed at 3 $^{7}/_{12}$ years of age, her mother was concerned about a mild waddling gait, but examination was normal. By 5 years of age, she had progressive knee valgus, internal tibial torsion, mild limitation of ankle dorsiflexion, and poor heel-toe transition, but was otherwise normal. Between 7 and 8 ½ years of age, neck stiffness, minimal limitation of elbow supination, and minimal wrist stiffness were evident, as were confounding features caused by an unrelated subtrochanteric femur fracture.

By age 11 ½ years, she had developed worsening neck stiffness; pain in her neck, hips, and knees; increased waddling; and decreased walking endurance. She had decreased mobility in her neck, elbows, wrists, hips, knees, and ankles. Proximal interphalangeal hand joints were prominent, and feet were normal except for incidental mild hallux valgus. Height was 135.4 cm (5th centile). Between 12-15 years of age,

kyphosis developed, mild scoliosis progressed, and joint stiffness and pain worsened, particularly in her hips. She is now a young adult and attends college. Joint immobility, particularly of the hips and knees, has continued to progress. She has severe pain with ambulation and can only walk for short distances with crutches. Both hip replacement and spinal fusion surgeries are anticipated.

Radiographs at 5 and 8 years of age were normal. By 12 years of age, her spine showed mild to moderate generalized platyspondyly, mild anterior wedging, loss of anterior substance and mild excess concavity of the vertebral bodies, and a lumbar scoliosis (Fig. 1d). There was mild, generalized joint space narrowing and metaphyseal irregularity (distal radii, distal ulnae, and distal femora; Fig. 1e-g), and phalangeal metaphyses of the hands were mildly prominent (Fig. 1g). Femoral heads were flattened and irregular, and femoral necks were short (Fig. 1e). The left femur was deformed secondary to previous injury.

The proband of Family 2 (Table 1) is a Korean boy who presented at age 6 years because of awkward gait and running difficulty. He was the product of a normal full-term pregnancy of unaffected, non-consanguineous parents and born with a weight of 3.7 kg. Motor and cognitive developments were normal, and he started to walk alone at 14 months of age. His parents noticed his slow, waddling gait and inability to run at 2-3 years of age. He was examined by a pediatric neurologist but not found to have neurologic or muscle disease. His height was 118.3 cm (z=+0.6, 73th centile), and weight 18.4 kg. His older brother and sister were unaffected. At age 8 $^{1}/_{2}$ years, he had an awkward gait and complained of difficulty squatting. He could run only slowly, and could barely climb stairs without the railing. Physical assessment revealed standing

posture with hips and knees flexed. He showed a wide-based gait with smooth heel-toe progression. The knees showed 20° flexion contractures. The hips and elbows showed mild limitation of motion. His height was 131.5 cm (z=+0.45, 67th centile) and arm span was 137 cm.

At 9 $^{4}/_{12}$ years of age, his knee and hip joint pain worsened, precluding sports activities. He had a stiff gait with hip abduction. The hip joints showed 15° flexion contractures, and further flexion was painful and limited to 100°. Nonsteroidal anti-inflammatories and diverse physical treatments lessened his joint pain. By 10 $^{1}/_{2}$ years of age, he could walk only with assistance due to joint pain and limb muscle weakness from prolonged disuse. Bearing weight on his legs resulted in hip, knee, and ankle pain. Motion at the hip joint was painful, with flexion contractures of 20° and flexion limited to 120°. The knee joint showed 20° flexion contractures.

His initial hip radiographs at 4 years of age (Fig. 1h) showed irregular contour of the femoral capital epiphyses, shallow acetabulum, and wide femoral neck with mild coxa valga. Skeletal survey at age 8 $^{1}/_{2}$ years (Fig. 1i-m) revealed generalized platyspondyly with wedged vertebral bodies at the thoracolumbar junction. His pelvis showed mild capital femoral epiphyseal dysplasia, coxa valga, and small, elongated ilia. Hand radiographs showed no distinct metaphyseal widening of the phalanges or metacarpals. Feet lateral radiographs showed large os trigonum bilaterally, and accentuated pes cavus. We prioritized a diagnosis of SED, Stanescu type and also included a broad category of unclassified spondyloepiphyseal dysplasias, possibly including a type II collagenopathy. Based on the proband's average to tall stature,

apparent spondyloepiphyseal dysplasia with coxa valga, and absent metaphyseal widening of the phalanges, there was no initial suspicion of PPRD.

Results

After multiple assessments failed to detect *WISP3* variants in Family 1, we performed WES on the proband, her unaffected mother, and her affected daughter. WES analysis and filtering in Family 1 (Supp. Materials and Methods) yielded 56 heterozygous variants, including a novel *COL2A1* variant (p.Gly207Arg). Sanger sequencing confirmed that this variant was present in the 2 affected individuals and absent in the unaffected mother of the proband (Supp. Figure S1), but we were unable to confirm *de novo* occurrence in the proband since paternal DNA was unavailable. In the proband of Family 2, we directly evaluated *COL2A1* for sequence and copy number variants because we suspected a type II collagenopathy (Supp. Materials and Methods). We detected the same heterozygous variant in the proband by full Sanger sequencing of *COL2A1* (Supp. Table S1) and confirmed *de novo* occurrence by Sanger sequencing of the mutation site in his unaffected parents (Supp. Figure S2). Functional prediction software indicates that the variant occurs at a conserved residue and is predicted to be damaging (Supp. Materials and Methods). This variant was submitted to a *COL2A1* locus-specific database at http://databases.lovd.nl/shared/variants/COL2A1, DB-ID COL2A1_000405, in three separate entries representing each individual with the variant.

Discussion

Interestingly, the clinical and radiological findings manifest somewhat differently between affected individuals in Families 1 and 2, although they share the same variant. The proband of Family 2 exhibited normal height until late childhood, and the proximal femora were broad and elongated with coxa valga. The proband of Family 2 also lacked enlarged phalangeal epimetaphyses of the hands, which is one of the hallmarks of PPRD and was prominent in Family 1.

COL2A1 encodes the alpha-1 chain of type II collagen, the primary collagen in articular cartilage. Many bone dysplasias are type II collagenopathies (MIM 120140), ranging from lethal to very mild and including Czech dysplasia. Although the individuals described here share many phenotypic features and the autosomal dominant mode of inheritance of Czech dysplasia, they lack major hallmarks including metatarsal shortening, which is thought to be uniformly characteristic of Czech dysplasia (Table 1). In addition, a single heterozygous *COL2A1* variant (p.Arg275Cys) has been ascribed to Czech dysplasia in all published cases (MIM 609162), while the individuals described here are heterozygous for a p.Gly207Arg variant.

Although there are few reports of Stanescu dysplasia, all describe a chondrocytic phenotype of PAS-positive, amylase-resistant cytoplasmic inclusions (Stanescu et al., 1984; Nishimura et al., 1998). Since we did not have access to chondrocytes from our patients, we were unable to test for this feature. However, given the phenotypic overlap with our cases, it is likely that Stanescu dysplasia is also a type II collagenopathy, and conceivable that it arises from the same mutation described here. We believe it would be

worthwhile to investigate the possibility of *COL2A1* mutation in other cases with this phenotype.

Rukavina et al. (2014) implicated a c.611G>T, p.Gly204Val *COL2A1* variant in a case of early-onset progressive osteoarthritis with mild spondyloepiphyseal dysplasia (Table 1; Rukavina et al., 2014). This patient shares many clinical features of the patients described here, suggesting similar consequences of the p.Gly204Val and p.Gly207Arg *COL2A1* variants.

Previous connections have been made between WISP3 and type II collagen. Overexpression of wild-type *WISP3*, but not PPRD variants, increases type II collagen in human chondrocytes (Sen, Cheng, Goldring, Lotz, & Carson, 2004; M. Wang et al., 2013). WISP3 normally co-localizes with type II collagen in the midzone region of chondrocytes, but WISP3 localization is disrupted by PPRD variants (Sen et al., 2004). Type II collagen secretion into the extracellular matrix also decreases in human chondrocytes transfected with mutant as compared to wild-type *WISP3* (M. Wang et al., 2013) We hypothesize that the *WISP3* variants found in PPRD and the *COL2A1* variant described here both decrease type II collagen, potentially disrupting the extracellular matrix and common downstream signaling pathways (Barber, Esteban-Pretel, Marín, & Timoneda, 2014).

These individuals heterozygous for the p.Gly207Arg *COL2A1* variant have a phenotype overlapping PPRD, SED, Stanescu type, and Czech dysplasia, although the third is clinically distinguishable by metatarsal shortening. The individuals described here have a disorder that further expands the phenotypic spectrum of type II collagenopathies.

We initially matched the common variant in these families at an academic conference. The team working with the proband of Family 2 described the variant in an abstract, which was subsequently recognized by the group following Family 1. While this story highlights the power of case matching, the probability that such matches will be made is enhanced by tools such as GeneMatcher (Sobreira, Schiettecatte, Boehm, Valle, & Hamosh, 2015). Based on our experience, we strongly support wider collaboration and data sharing, particularly for researchers and clinicians following rare disorders.

Acknowledgements

We are grateful to our patients and their families for participation and to Dr. James Hyland of Connective Tissue Gene Tests for molecular analysis of *WISP3* in Family 1. Our work was supported in part by grants from the US NIH/NHGRI (T32GM07814; 1U54HG006542) and the Genome Technology to Business Translation Program of the National Research Foundation funded by the Ministry of Science, ICT & Future Planning of the Korean Government (NRF-2014M3C9A2064684).

Chapter 2 Figures

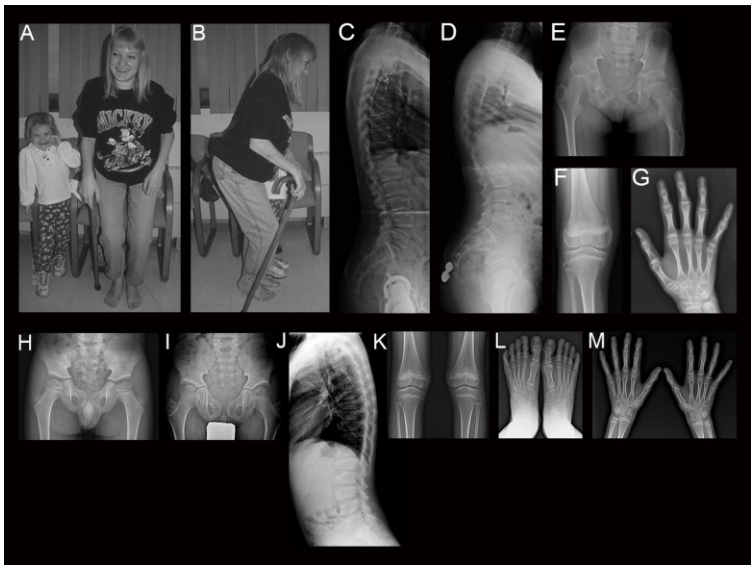


Figure 1. Clinical features of patients. a-c: Proband of Family 1 at 23 years. Note marked joint contractures resulting in Z posture. **c:** Lateral spine radiograph shows platyspondyly and irregular degenerative changes. **d-g**: Radiographs of the daughter of the proband of Family 1 at 12 years. **d:** Mild anterior wedging and platyspondyly of lateral spine. **e:** Hip and pelvic features include flattened and irregular femoral heads, short femoral necks, and deformed shaft of the left femur secondary to prior trauma. **f:** Knee radiograph shows joint space narrowing and metaphyseal irregularity. **g:** Note metaphyseal prominence and joint space narrowing at distal radius and distal ulna. **h-m:** Radiographs of the proband of Family 2 at 4 years (**h**) and 8 years (**i-m**). **h:** Note irregularity and mild flattening at articular surfaces of capital femoral epiphyses. Acetabulum is flat, and femoral necks are broad with mild coxa valga. **i:** Capital femoral epiphyses show mild flattening and dysplastic trabeculation. Femoral necks are broad and elongated with further coxa valga. Ilia are somewhat narrowed and elongated. The

acetabulum shows deepening with irregularity of its contour. **j:** Note moderate spinal platyspondyly with anterior wedging and irregularity of the endplates. **k:** Knee joints show mild flaring of the distal femora and proximal tibiae. Epiphyses are normal. **l-m:** Hands and feet are normal.

Chapter 2 Tables: Table 1. Comparison of Clinical and Molecular Features of Relevant Disorders

	Family 1	**Family 2**	**Rukavina et al. [2014]**	**Czech Dysplasia**	**SED, Stanescu type**	**Progressive Pseudorheumatoid Dysplasia (PPRD)**
Mode of Inheritance	Autosomal dominant	*De novo*	Autosomal dominant	Autosomal dominant	Autosomal dominant	Autosomal recessive
Family Origin	U.S.A.	Korea	Unknown	Multiethnic	Multiethnic	Multiethnic
Major Clinical Features						
Short Stature	+	-	+	-	-	+
3rd and 4th Toe Shortening	-	-	-	+*	-**	-
Hearing Loss	-	-	-	+/-	+/-***	-
Joint Pain	+	+	+	+	+	+
Progressive Joint Immobility	+	+	+	+	+	+
Joint Replacement Surgery Needed	+	Not yet	+	+	+	+
Major Radiologic Features						
Progressive Platyspondyly	+	+	+	+	+	+
Progressive Joint Space Narrowing	+	?	+	+	+	+
Metaphyseal Enlargement, Hands	+	?	+	+/-	+	+
Metatarsal Shortening	-	-	-	+*	-	-
Mutated Gene	*COL2A1*	*COL2A1*	*COL2A1*	*COL2A1*	Unknown	*WISP3*
Genetic Variant****	c.619G>A (heterozygous)	c.619G>A (heterozygous)	c.611G>T (heterozygous)	c.823C>T (heterozygous)	N/A	Allelic Heterogeneity [Hurvitz et al., 1999; Delague et al., 2005; Zhou et al., 2007; Yue et al., 2009; Temiz et al., 2011; Dalal et al., 2012; Garcia Segarra et al., 2012; Sun et al., 2012]
Protein Variant****	p.Gly207Arg (heterozygous)	p.Gly207Arg (heterozygous)	p.Gly204Val (heterozygous)	p.Arg275Cys (heterozygous)	N/A	Allelic Heterogeneity (See references above)
SIFT/ PolyPhen-2 Scores	0/1	0/1	0/0.999	0/0.999	N/A	N/A

*Although the patients described by Burrage et al. [2013] did not have disproportionate shortening of the 3rd and 4th toes, they did have radiologic evidence of generalized shortening of the metatarsals.

**One proband described by Nishimura et al. [1998] had shortening of his 2nd and 3rd toes and 2nd fingers; this is distinct from the 3rd and 4th toe shortening observed in Czech dysplasia.

***The original proband described by Stanescu et al. (1984) had sensorineural hearing loss.

**** *COL2A1* variant positions are based on RefSeq transcript NM_001844.4.

Chapter 2 Supplement

Variant Detection in Family 1

WISP3 was initially investigated in the proband of Family 1 by Hurvitz et al. (1999). Subsequent clinical molecular analysis (courtesy of James Hyland, Connective Tissue Gene Tests) failed to demonstrate pathogenic *WISP3* variants in either the proband or her daughter. All five exons of *WISP3* were PCR-amplified and sequenced using the ABI 3730 sequencer, and *WISP3* was analyzed for copy number variants using a CTGT high-density targeted array. Both these methods failed to detect any pathogenic sequence or copy number variants in *WISP3*.

Next, we performed whole exome sequencing on the proband, her unaffected mother, and her affected daughter at the Baylor-Hopkins Center for Mendelian Genomics. The patients were counseled on the possible outcomes of whole exome sequencing and signed a consent form approved by the Johns Hopkins University School of Medicine Institutional Review Board. To capture the CCDS coding regions and adjacent intronic sequences, we used the Agilent SureSelect Human All Exon V4 51Mb capture kit to obtain paired-end 100 bp reads on the Illumina HiSeq2500 platform. Reads were aligned to the reference genome (NCBI human genome assembly build 37; Ensembl core database release 50_361 (Hubbard et al., 2009)) using the Burrows-Wheeler Alignment (BWA) tool (H. Li & Durbin, 2009), and single nucleotide variants (SNVs) and small insertion-deletions (indels) were identified using SAMtools (H. Li et al., 2009). We performed local realignment and base call quality recalibration using GATK2.3.9 (DePristo et al., 2011; McKenna et al., 2010). We then selected for variants with a minimum of 8 reads, root mean square mapping quality of 25, strand bias p-value below

10^{-4}, and end distance bias below 10^{-4}. Using a presumed autosomal dominant inheritance model, we prioritized variants using the Analysis Tool of PhenoDB (Hamosh et al., 2013). We selected for rare heterozygous variants with MAF<0.01 in the 1000 Genomes Project (Build 20130502; 1000 Genomes Project Consortium et al., 2012); dbSNP builds 126, 129, and 131 (Sherry et al., 2001); Exome Variant Server (release ESP6500SI-V2); and our in-house control database (CIDRVar 51Mb), and nonsynonymous exonic or splicing SNVs and indels which were present in affected individuals but absent in unaffected individual.

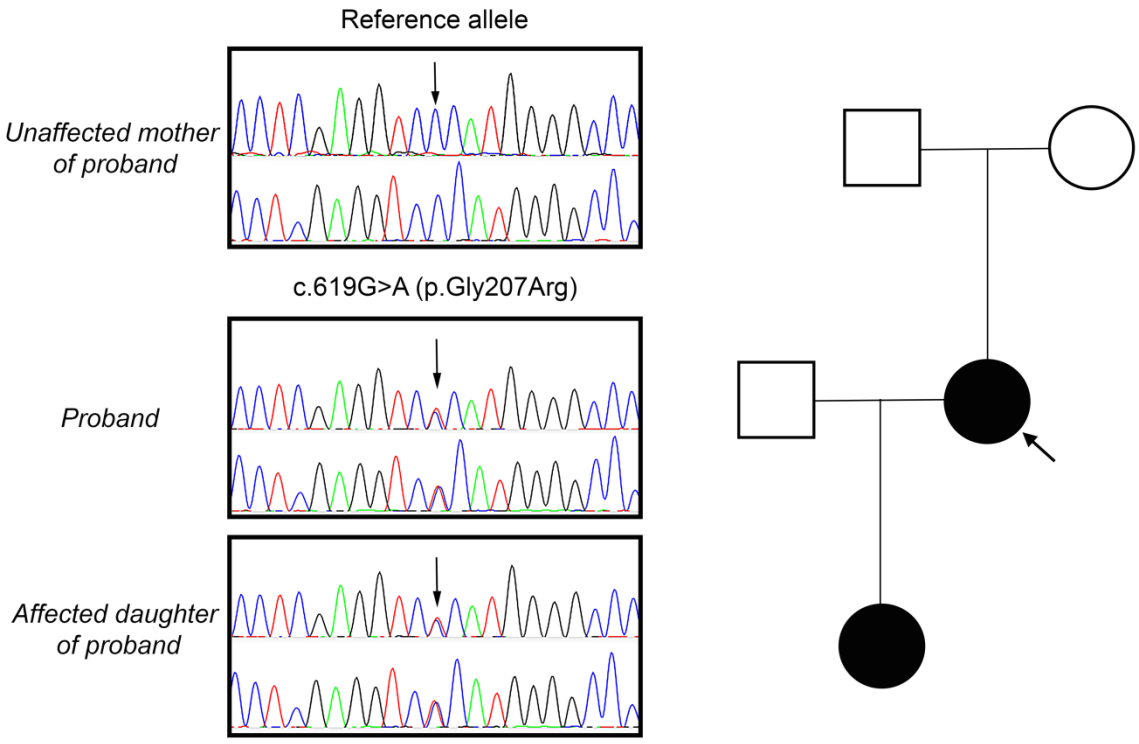
Variant Detection in Family 2

We directly sequenced the *COL2A1* gene in the proband of Family 2 because his radiological findings suggested a spondyloepiphyseal dysplasia. Genomic DNA was extracted from peripheral blood leukocytes. The *COL2A1* gene was screened by direct Sanger sequencing of all coding exons and their flanking regions (Supp. Table S1). Multiplex ligation-dependent probe amplification (MLPA) analysis was performed to screen gross gene deletions and duplications involving this gene using the SALSA P214 *COL2A1* (MRC-Holland, Amsterdam, Netherlands). Upon detection of the variant of interest, Sanger sequencing of the mutation site was performed on both unaffected parents of the proband to confirm co-segregation of the variant with the phenotype (Supp. Figure S2).

Prediction of Variant Pathogenicity

The novel *COL2A1* variant (RefSeq transcript NM_001844.4: c.619G>A, p.Gly207Arg) was assessed for putative conservation and functionality using a variety of established methods. The variant has a GERP++ rejected substitutions score (Genomic

Evolutionary Rate Profiling score, GERP++ version) of 4.81 (Davydov et al., 2010); a SIFT score of 0 (SIFT v. 1.03, score range 0-1.0, with scores ≤0.05 predicted damaging; Kumar, Henikoff, & Ng, 2009); and a PolyPhen-2 score of 1 (score range 0 (benign) to 1.0 (probably damaging), PolyPhen-2 v.2.2.2; Adzhubei et al., 2010). These scores indicate that the variant occurs at an evolutionarily conserved residue and is predicted to be damaging.

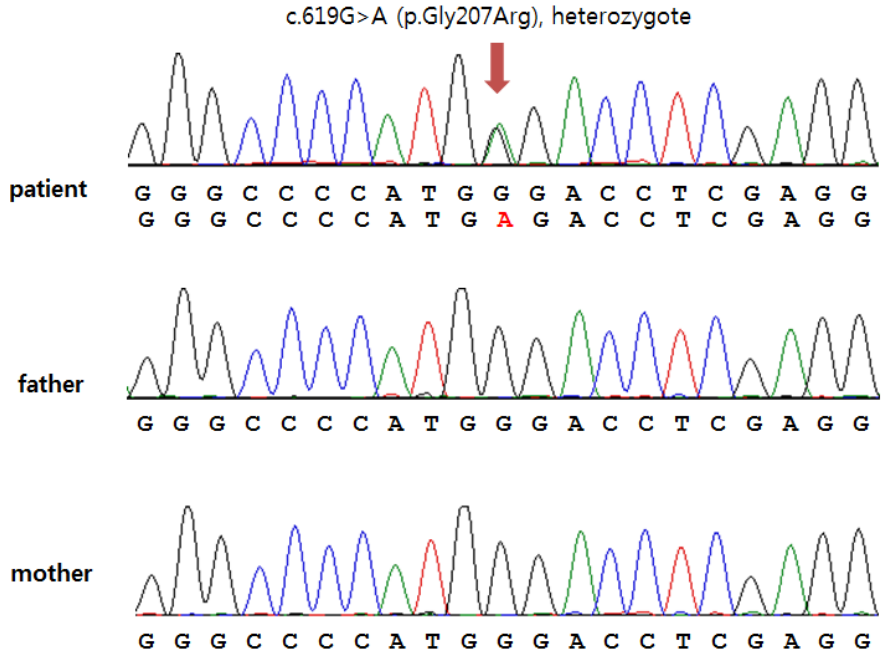


Supp. Figure 1. Sanger validation of *COL2A1* variant in Family 1. Sanger sequencing was performed to validate co-segregation of the *COL2A1* variant c.619G>A

(p.Gly207Arg, RefSeq transcript NM_001844.4) with the phenotype in Family 1. Whole exome sequencing originally detected this variant in the proband and her affected daughter, but not in her unaffected mother. Primers were designed for amplification and subsequent Sanger sequencing of a 357 bp fragment encompassing the *COL2A1* variant. Results confirm co-segregation of the *COL2A1* variant with the phenotype in Family 1 (the variant is present in the proband and her affected daughter, but absent in her unaffected mother). Primer sequences were as follows:

Forward CACCAGGATTGCCTTGAAAT

Reverse TAGAGTGTTGCGTGGGGAGT



Supp. Figure 2. Sequence analysis of *COL2A1* exon 9 in Family 2. The proband was heterozygous for the missense variant c.619G>A (p.Gly207Arg, RefSeq transcript NM_001844.4). This missense variant was not detected in his unaffected parents, indicating that it arose *de novo* in the proband.

Supp. Table 1. Primers used for Sanger sequencing of *COL2A1* gene in the proband of Family 2.

Exon	Forward primer sequence	Reverse primer sequence
1	AGCGTGACTCCCAGAGAGG	GGTGAACTTCTGCGTGTCCT
2	CCCAGCCTACATTCTTCAGC	GTGGCCTTTCCTTTCTACCC
3-6	AGAGGGGTTCAGTGGAGTCA	AGGTGACCAGGCCCTTAAAC
7	AAAAGCAATGCTGCTTGACC	GCCCTTAGCACCACAGTCTC
8	ACTTTCCAGGCATCAGTGGT	CTGGACAGCAGCCATGTTTA
9-10	CTGTCCAGGAGAAAGCAAGG	GGGGGCTGTTAGGACAGAGT
11	CCACCAGGGACTCTGTCCTA	ACTTGCCAGCACAGTCACAG
12	GGTGAGGACAGAGGAGCTTG	TCCCTCATGAGAGCCTGAGT
13-14	AGGGGTGGTCTGGGTTTTAC	CAGGATGTAGGGCTGGTGTT
15	GCATTGTCCCTCACAGGACT	ACAAGCACACCTTGCTCCTT
16	TCTGAGGCATCCTGGCTAGT	CACACTGGGGGTGAATTCTT
17	AGCCAGACTCCGTCTCAAAA	CCTCCCCCTTTCCAGTAGAC
18	CTGAAGCCAGGCAAAGTTTC	ACTGCGAGTGTCTGGCTTTT
19	TGGGTGCATGTGCATAATTT	GGCATAGGTGCTGTCCATTT
20-21	ATGTCCCTGAAATGGACAGC	AAGAGAGGCAGGTCCTCACA
22	CTCTTTCCCATGCTCTCCTG	AGGGAAGTGAGAGGGGCTAA
23	CAGTTGCCAAGGCTACTTCC	CATGTCAGTCTGGTGGTTGG
24-26	ACAGCTACTGCTCCCCAGAA	GCCTACCATCTACCCCCTGT
27	AAGCTCAGCGGTGTCTTGTT	CTTCTGTTCAGGGGCCATAA
28	TTATGGCCCCTGAACAGAAG	TGACCAATGGCAAACAGAAG
29	GTACCGTGGAGGTCTGGAAA	CTTCCTGCACCACTCAGACA
30-31	TGTCTGAGTGGTGCAGGAAG	GACATGGGTCCAGGACATTC
32-33	AGCCTCCACAGATGACACAA	GACAGCACCTCTCCCTGAAC
34	TGCATCCTGTAGGGGTCTTC	TTCATCACCAGGTGCCATAA
35-36	ACATCCTTTCCCCTGTTCCT	ACAAGACAGAACCGCCTTTG
37	ATGGGCTGCACAGTAACACA	CTCTCCTCCCACCCTAGACC
38-39	CCATCATGTGTTCAGGGTGA	CTACCAACATGGGGGTGTTC
40	CACAAGTCTCCTGTGGCAGA	TCCACAGTCAGCACTTCAGG
41	GTTCAGCTCCCTGAAGTGCT	GGTCCCCATGATCAGTTAGC
42	AGGTAGAGAAGCCCCCAGAG	TGAGAGAGGAGAGGCAGGAG
43-44	TCTGAGCTCACAGAGCATGG	CAAGTTTCCCTCCTCCTTCC
45-46	GAAGGAGGAGGGAAACTTGG	CGTCAGAGAAAAGCCAGGAC
47-48	TCAGTAAAAGCCGCCTTCC	TGGGGGAACAGCTTTATGTC

49	GACAGCTTGGGATCACCCTA	AGCGAGAGGCTGATTCATGT
50	ATGAATCAGCCTCTCGCTGT	CAGGCTGATGCCCTAAAAGA
51	CTGACTCCTTCCCTGCTGTC	CCCAGCTCTGCCCTGTACTA
52	CCTCATCCCTTGTCCCTGTA	CAGCAGGGAGCTAGTTGGAC
53	GCAGAGATGGCTCCTCAAAC	TCATGCCTCTGATGATCCTG
54	ACAGGTTGGGAGCTCACTGT	TCTGCCCAGTTCAGGTCTCT

Chapter 3. Assessment of incidental findings in 232 whole exome sequences from the Baylor-Hopkins Center for Mendelian Genomics

Chapter 3 is reprinted from the final accepted journal article:

Jurgens J, Ling H, Hetrick K, Pugh E, Schiettecatte F, Doheny K, Hamosh A, Avramopoulos D, Valle D, Sobreira N. (2015). Assessment of incidental findings in 232 whole exome sequences from the Baylor-Hopkins Center for Mendelian Genomics. *Genet Med*. 17(10): 782-8.

Julie Jurgens's Contributions: Analysis of incidental findings in 232 exome sequences; conceptualizations of portions of incidental findings analysis (with Dr. Nara Sobreira); conceptualization and drafting of manuscript and figures; coordination of revisions from other co-authors; coordination of submission process; drafting of response to reviewer comments and integration of additional responses from other co-authors; manuscript finalization with journal for publication

Introduction

Incidental findings in whole exome and whole genome sequencing are variants of known or possible pathology identified in genes unrelated to the initial reason for which sequencing tests were ordered. As large-scale sequencing is increasingly used in a clinical setting, standardizing procedures for identifying, classifying, and reporting variants is of great importance. In 2013, the American College of Medical Genetics and Genomics (ACMG) released a recommendation suggesting that known pathogenic or expected pathogenic constitutional variants detected in 56 genes representing 24 conditions with well-known guidelines for prevention or treatment should be reported to patients sequenced in a clinical setting (Green et al., 2013). Shortly after the release of

these recommendations, the Clinical Sequencing Exploratory Research (CSER) Consortium and Electronic Medical Records and Genomics (eMERGE) Network followed by releasing a set of guidelines for the return of incidental findings within genomics research studies (Jarvik et al., 2014). These guidelines recommended that medically actionable incidental variants be returned to willing research participants, but exempted researchers from an obligation to search for these variants.

Although incidental variants are unexpected, when variant pathogenicity is certain, the information enables patients to make proactive decisions about their health and inform their physician. However, some variants are of uncertain significance, or initial interpretation of variant pathogenicity may be subject to change (Bell et al., 2011). Reporting these variants of uncertain significance (VUS) to patients may lead to misinterpretation of the results, produce unnecessary anxiety, and diminish confidence in genetic test results in general. Thus, transmittal of these results requires accurate and thoughtful genetic counseling and careful evaluation of available sources of variant classification to maximize benefits and minimize harm to patients.

At the Baylor-Hopkins Center for Mendelian Genomics (BHCMG, http://mendeliangenomics.org), we utilize whole exome sequencing to detect variants responsible for Mendelian disorders with unknown molecular bases. Since we sequence across the exome, we not only capture variants that are responsible for the patient phenotype, but also incidental variants which may predispose and/or cause other unrelated disorders. Returning meaningful results to research participants through their physicians is a goal of our center, so understanding the clinical significance of incidental

findings detected at the research level and knowing which of these should be returned to patients is of special interest.

Although the ACMG guidelines for reporting incidental findings are geared toward a clinical setting, they provide a well-defined, targeted set of genes from among the many interrogated in research sequencing. Thus, our results on this set may inform policies for handling incidental findings. Here we examine the implementation of the ACMG clinical guidelines within a research setting by assessing the spectrum of rare, functional variation in the 56 genes in whole exome sequences of 232 individuals in 89 families sequenced at the Baylor-Hopkins Center for Mendelian Genomics for a variety of Mendelian disorders with unknown molecular bases, some of which have features that overlap with the disorders covered by the ACMG list. We then assessed variant interpretation by evaluating classification and reportability according to available databases and approaches taken by others in the literature.

Materials and Methods

All research participants were counseled on the possible outcomes of whole exome sequencing and signed a consent form approved by the Johns Hopkins University School of Medicine Institutional Review Board. Participants chose whether they would like to receive information regarding primary or incidental variants found in the course of this study and were given the opportunity to opt out or in at any time.

Whole exome sequences from 232 individuals in 89 families sequenced at Johns Hopkins University (JHU) were analyzed for rare splicing or nonsynonymous SNVs and

indels in the 56 reportable genes from the ACMG guidelines using the analysis feature of the web-based system PhenoDB (Hamosh et al., 2013).

Briefly, whole exome sequencing was done on the Illumina HiSeq2000 platform (Illumina, Inc., San Diego, CA) using paired-end 75 or 100 bp reads. 51 Mb of genomic DNA comprised of CCDS exons and adjacent intron sequences were captured with Agilent SureSelect Human All Exon V4 51Mb Kit (Agilent Technologies, Santa Clara, CA). FASTQ files were aligned to the reference genome (GRCh37; Ensembl core database release 50_361; Hubbard et al., 2009) with the Burrows-Wheeler Alignment (BWA 0.5.10) tool (H. Li & Durbin, 2009) resulting in SAM/BAM output (H. Li et al., 2009). PCR duplicates were flagged with Picard (http://picard.sourceforge.net). Local realignment around indels, base call quality score recalibration and reduced reads BAMs were performed with GATK 2.3-9 (McKenna et al., 2010). Multi-sample SNV and indel calling was performed with GATK's UnifiedGenotyper on the reduced reads BAM files. Variant sites were filtered with GATK's Variant Quality Score Recalibration best practices (DePristo et al., 2011), and heterozygous genotypes were excluded if they did not have at least 5 alternate allele reads. Sequencing coverage of the 56 genes is listed in Supplementary Table S1.

We selected for nonsynonymous SNVs, exonic indels, and splice junction variants (2 most 5' and 3' intronic bases) in each of the 56 reportable genes. Nonsynonymous SNVs and exonic variants were defined by RefSeq Gene coordinates (Pruitt, Tatusova, & Maglott, 2004). Next, we filtered these variants on the basis of their minor allele frequencies (MAF), with exclusion of variants with MAF $\geq$ 0.01 in the 1000 Genomes Project (April 2012 release; 1000 Genomes Project Consortium et al., 2012), Exome

Variant Server (http://evs.gs.washington.edu/EVS/), dbSNP build 137 (Sherry et al., 2001), and our in-house control database (CIDRVar 51Mb), which includes data from 50 individuals sequenced in-house with the Agilent51Mb exome capture kit (Agilent Technologies, Santa Clara, CA).

Next, we analyzed the variants for prior classification in the Human Gene Mutation Database (HGMD; Stenson et al., 2003), ClinVar (Landrum et al., 2014), and the *Emory* Genetics Laboratory *Variant* Classification Catalog (Bean, Tinker, da Silva, & Hegde, 2013). We then defined the burden of variants in each individual sequenced and the mutability of each gene (number of rare functional variants/gene and number of rare functional variants/kb of coding region), and the classification and reportability of each variant.

Results

In the 232 exomes, we identified 249 distinct variants in the 56 gene ACMG target set, some of which were seen in more than one individual, for a total of 391 variants. A total of 124 variants were shared between family members, and 18 variants were shared between unrelated individuals. There was an average of 1.69 variants per individual, with a range of 0-7 variants per individual. There were only 45 people (19.4%) with no variants, and in the 232 exomes we found at least one variant in 45 of the 56 genes (80.4%; Supplementary Table S1).

We then stratified these 249 distinct variants by type (Table 1). Most variants were missense (231/249, or 92.8%). There were also 4 nonsynonymous/splice variants

(1.6%), 2 splice variants (0.8%), 3 frameshifting indels (1.2%), 5 nonframeshifting indels (2.0%), and 4 nonsense variants (1.6%).

Next, we checked the classification of these variants in three databases: HGMD, ClinVar, and Emory. Of the 249 distinct variants identified, 126 (50.6%) were classified by at least one of the three databases: 54/126 (42.8%) were represented solely in HGMD; 44/126 (34.9%) were in HGMD and ClinVar; 23/126 (18.3%) were in ClinVar alone; 2/126 (1.6%) were in Emory alone; and another 3/126 (2.4%) were in all three databases (Figure 1).

In total, 101/249 variants (40.5%) were listed in HGMD (94 missense SNVs (93.1%), 2 nonsynonymous exonic/splicing variants (2.0%), 2 splicing variants (2.0%), 1 nonsense SNV (1.0%), and 2 nonframeshifting indels (2.0%)). More importantly, HGMD classified 74 (72.8%) as disease-causing mutations (DM), 24 (24.3%) as possible disease-causing mutations (DM?), 1 (1.0%) as a functional polymorphism (FP), and 2 (1.9%) as disease-associated polymorphisms (DP) reported to be in significant association with disease ($p<0.05$) along with some evidence of functionality.

Of the 148 variants not classified by HGMD, 23 variants in 13 genes were described by ClinVar and 2 variants in 2 genes were listed in Emory. This left a total of 123/249 variants (49.4%) that were not present in any database (Figure 1).

The classification of variants within these three databases was often discordant. Of the 101 variants represented in HGMD, 48 were also in ClinVar. Three of these 48 variants were also in Emory. Almost half of these shared variants (22/48 variants, or 45.8%) were given discordant classifications among databases (Supplementary Table S2).

Next, we examined the effects of gene size and evolutionary constraint on the number of functional variants within each gene, in a manner similar to that described by Petrovski et al. (Petrovski, Wang, Heinzen, Allen, & Goldstein, 2013). Some genes consistently had greater variation than others. For instance, *BRCA2*, *APOB*, *CACNA1S*, and *DSP* had the greatest number of variants/gene (Supplementary Table S1). Since increased coding length increases the target size for mutation, larger genes are expected to have a higher number of variants than smaller genes. To explore this possibility, we analyzed the number of variants per gene against the size (kb) of the consensus coding sequence (CCDS) of the gene. We found that the number of variants per gene generally increased with the size of coding sequence, and the number of variants showed a 62.8% correlation with the size of the consensus coding sequence (Figure 2).

We also found that some genes (e.g. *RB1*, *PMS2*, and *MLH1*) consistently have a lower variant density (variants/kb coding sequence) than others (e.g. *MYBPC3*, *TMEM43*, and *MYL3*) (Supplementary Table S1). The values of variant density range from 0.4 variants per kb in *RB1* to 3.4 variants per kb in *MYL3*. As described by others, a possible explanation for this variation in mutational burden despite correction for coding sequence length is varying degrees of evolutionary constraint (Petrovski et al., 2013). Certain genes are under a higher degree of purifying selection than others, and this is likely reflected by less variation per unit coding length.

Determining Reportable Variants

To determine which variants should be reported to our research participants, we adopted a stringent set of measures outlined in a recent paper (Dorschner et al., 2013). Their criteria incorporates various lines of evidence in assessing variant pathogenicity,

including comparison between variant minor allele frequencies and incidences of the corresponding disorders; familial segregation data; observation in unrelated affected individuals; *de novo* events in a trio; and protein truncation in disorders caused by haploinsufficiency. If a variant is classified as pathogenic or likely pathogenic by these criteria, it is considered reportable back to the patient.

We applied these criteria to all 391 variants in the 56 ACMG genes detected in 232 exomes. In total, we found 2 pathogenic variants (in *MSH2* and *MYLK*) and 3 likely pathogenic variants (in *LMNA, MYBPC3,* and *MUTYH*), as shown in Table 2. One of these genes, *MUTYH*, is responsible for autosomal recessive conditions (*MYH*-associated polyposis; adenomas, multiple colorectal, *FAP* type 2; colorectal adenomatous polyposis, autosomal recessive, with pilomatricomas; MIM 608456 and 132600). Since current ACMG guidelines recommend reporting only individuals with homozygous variants in *MUTYH* (Green et al., 2013), we would not report this heterozygous variant to the patient. Since the patients in our cohort underwent whole exome sequencing to explain a variety of potential Mendelian disorders, some of which have overlapping features with the disorders covered by the ACMG list, we recognized that our cohort may have an enrichment of reportable variants as compared to the general population. Indeed, two patients, one with a likely pathogenic variant in *LMNA* and one with a likely pathogenic variant in *MYBPC3,* had clinical phenotypes which included dilated cardiomyopathy, so we did not count these two cases toward our final number of incidental variants. In total, this means that 2/232 individuals, or 0.86% of our sample, had a reportable incidental variant in one of the 56 ACMG genes.

Discussion

In June of 2014, Jarvik et al. released recommendations for dealing with incidental variants in the research arena. The authors suggested that researchers be exempted from an obligation to search for variants outside the intended scope of their study, but that actionable variants be reported to research participants when discovered (Jarvik et al., 2014). The present study describes the workflow for identification and classification of incidental findings in the Johns Hopkins component of the Baylor-Hopkins Center for Mendelian Genomics based on the ACMG recommendations and on the classification criteria from Dorschner et al. (Dorschner et al., 2013).

In 232 exomes, we identified a total of 391 rare (MAF < 1%) variants in the 56 ACMG genes (1.69 variants per individual), 249 of which were distinct variants. Most of these variants were missense (231/249, or 92.8%), and half were not classified by any of the three databases (HGMD, ClinVar, and Emory) that we used to assess variant classification (123/249 variants, or 49.4%). Since these novel variants were not represented in the variant databases, they could not be utilized for comparative analyses of classifications among the three databases.

Next we used the criteria described by Dorschner et al. to classify the 391 total variants identified in 232 exomes and found that 2/232 individuals, or 0.86% of our sample, had an incidental reportable variant in one of the 56 ACMG genes. Overall, this analysis was quite time-consuming, but we were able to automate parts of it such as filtering variants whose frequency was higher than that predicted by the frequency of their associated disorders. This step reduced the number of distinct variants from 249 (obtained in our initial analyses using a 1% MAF cutoff) to 163, so only 65.5% of the

original variants required further assessment. We were also able to automate selection of variants in the HGMD, ClinVar, and Emory classification databases using our web-based system, PhenoDB (Hamosh et al., 2013). Although we did not consider the classifications from these databases in support of variant pathogenicity or reportability, the databases provided useful citations from the literature and clinical observations for some variants. This information reduced the time required for manual review. Protein truncation was automatically predicted, but required manual inspection to confirm that protein truncation is an established cause of the disorder associated with the respective gene. The most time-consuming part of the process was reviewing the literature for each variant to count familial segregations, *de novo* events, and singleton reports. This step required additional expertise regarding diagnoses of the disorders on the ACMG list. Moreover, evaluation of cases in the literature was somewhat subjective.

Several other studies have analyzed reportable incidental findings within whole-exome or whole-genome sequences to-date. The study by Dorschner et al., whose criteria we adopted for our analysis of reportable variants, found that 14/1000 individuals sequenced (1.4%) had a pathogenic or likely pathogenic reportable variant in one of the 52 ACMG genes responsible for adult-onset conditions (Dorschner et al., 2013). Johnston et al. found that 8/572 (1.39%) individuals sequenced through the ClinSeq project had a reportable variant in one of 37 cancer genes; 23 of these are on the ACMG list (Johnston et al., 2012). These numbers are similar to ours and to those in the original ACMG guidelines, which anticipated that 1% of individuals sequenced would have a reportable variant in one of these 56 genes (Green et al., 2013).

Another recent report by Lawrence et al. detected 27/543 individuals (~5%) sequenced through the NIH Undiagnosed Diseases Program with an incidental variant in one of the 56 ACMG genes (Lawrence et al., 2014). Their method differed from ours in a few ways which may have contributed to the discrepancy in frequency of incidental findings. For instance, they used reported functional studies as evidence of pathogenicity; required fewer informative meioses to count familial co-segregation events with disease; did not apply minor allele frequency cutoffs; and used designations in variant classification databases as evidence in support of variant pathogenicity. In addition, a few individuals in their cohort had a phenotype related to the reported incidental finding, and some incidental variants were repeated due to familial transmission (Li Wang et al., 2010). These differences in methodology may have contributed to the discrepancies in numbers of reportable incidental findings.

During this process we recognized that identifying and classifying the multitude of incidental variants that arise from next-generation sequencing is a time-consuming and subjective process dependent on the particular methodologies of each individual investigator. We confirmed this last observation when we found that almost half (22/48, or 45.8%) of the 101 variants represented in at least 2 databases were given discordant classifications between them. We compared these database classifications to our variant classifications from the Dorschner et al. method (Dorschner et al., 2013) and found that variant classifications from the ClinVar and Emory databases, but not the HGMD database, were generally consistent with our variant classifications (Supplemental Table S2). This is likely because HGMD aims to provide a comprehensive listing of all reported variants and draws its classifications from reports in the literature (Stenson et al.,

2003) that may not be accurate, consistent, up-to-date, or sufficient to provide support of variant pathogenicity. For these reasons, HGMD is generally less conservative than the other databases.

The abundance of rare and novel missense variants in our cohort made accurate assessment of variant pathogenicity more challenging, since all novel and many rare missense variants must initially be classified as variants of unknown significance (VUS). For many rare variants, familial segregation data or *de novo* events in a trio were not available or sufficient to define variants as pathogenic or likely pathogenic.

When we observed *de novo* variants in a trio or co-segregation of a variant with disease, family information could be used as evidence supporting pathogenicity and reportability. Although failure of a variant to segregate might be considered evidence against pathogenicity, we did not use this information in our pathogenicity evaluations, since many disorders on the ACMG list are adult-onset and could be late to manifest in unaffected family members. Family sequence data is also useful for determining the phase of multiple heterozygous variants in a single gene, and thus can be used for finding compound heterozygous variants. When individuals rather than families are sequenced, it may be more difficult to classify variants as pathogenic or likely pathogenic since these sources of information are not available. However, support for pathogenicity can be obtained in other ways, such as literature searches. Thus, family information is helpful but not necessary for determining variant pathogenicity in most cases.

We found that prediction is even more difficult in genes that are subject to more variation because of their increased size or tolerance for variation. Since functional

information was not available for most variants, 152/249 distinct variants (61.0%) were classified as VUS.

Some variants had been previously described in the dbSNP, EVS, and/or 1000 Genomes databases, and of 249 distinct variants, 126 (50.6%) were classified by at least one of the three databases used to assess variant classification (HGMD, ClinVar, and/or Emory). The available classifications in these cases were helpful but not always clear, mainly due to discrepancies among bioinformatic prediction databases or functional studies; low numbers of unrelated affected individuals with the same variant; misreporting of variant pathogenicity in the literature or classification databases; or occurrence of variants in large or highly mutable genes. These inconsistencies all point to the need for a single, centralized database and methodology for variant classification.

Our study has several limitations. First of all, many individuals in our cohort were ascertained for various Mendelian disorders. Because of this, two of the reportable findings could not be considered incidental, in that they may actually have contributed to the phenotype of the patients. In addition, our cohort was relatively small, and 73% of individuals self-identified with European descent. Moreover, we may have missed some variants due to low coverage in some regions of the 56 ACMG genes. Another limitation was the minor allele frequency cut-offs that we employed. For our early comparisons of variant classification between databases, we used a 1% MAF cut-off, which may be too high for some disorders and necessitate more time for manual review of variants. On the other hand, our later assessments of variant reportability used more stringent minor allele frequency cut-offs based on expected disease prevalences. This increased stringency may have resulted in exclusion of variants which were in fact pathogenic. In addition, the

method we used to find reportable variants fails to consider functional studies. Because of this, we may have excluded some variants which have been functionally characterized as pathogenic or included variants which show milder or less conclusive functional consequences.

An alternative sequencing method like whole-genome sequencing would increase detection of incidental findings as well as VUS, owing to the increased coverage of this method and the challenge of interpreting noncoding variants. Whole-genome sequencing will uncover structural variants and copy-number variants as well. However, the ACMG recommendations currently apply only to SNVs and indels, so this additional variation—some of which may be significant—would not be considered under the current guidelines (Green et al., 2013).

Based on this experience, we suggest that clinicians and researchers interested in returning incidental variants found by next-generation sequencing adopt a uniform, well-defined set of criteria for variant classification. Here we applied the ACMG recommendations and the classification criteria defined by Dorschner et al. (2013), but we suggest that the addition of bioinformatic prediction tools, genic intolerance scores, and functional studies to the criteria would make variant classification more complete. We also expect that research and clinical laboratories performing next-generation sequencing will soon be more engaged in submitting variants to manually curated classification databases, improving interpretation of variants and counseling of patients.

Acknowledgements

We would like to thank all of our research participants who made this study possible. J.J. is a trainee of the Predoctoral Training Program in Human Genetics at Johns Hopkins University and is supported by the NIH grant T32GM07814. This work was funded by a grant from the National Human Genome Research Institute (U54HG006542).

Chapter 3 Figures

Figure 1. Classification of the 249 variants by the HGMD, ClinVar, and Emory databases.

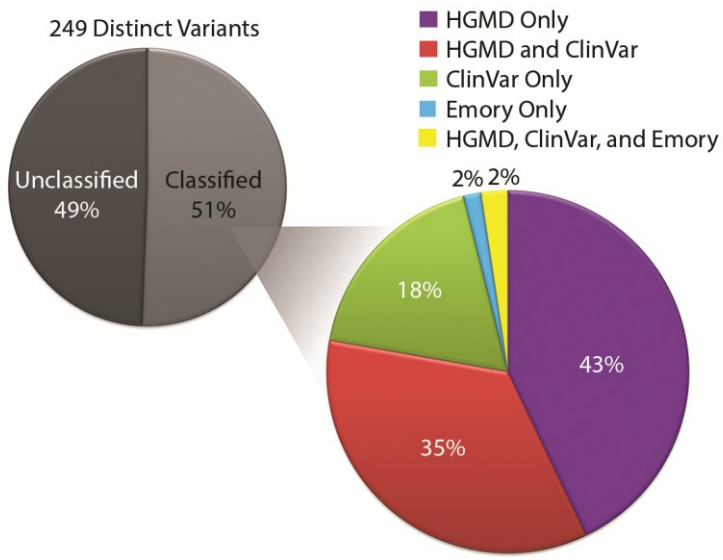
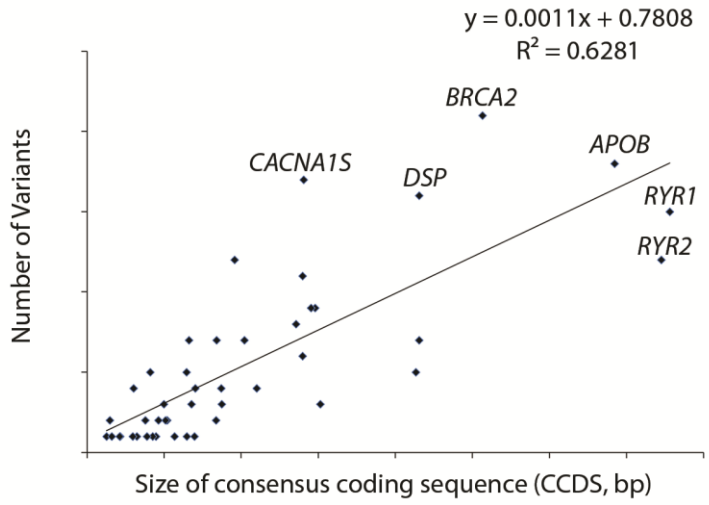


Figure 2. The number of variants per gene plotted against the size (kb) of the consensus coding sequence (CCDS) of the gene.



Chapter 3 Tables

Table 1. Classification of the 249 variants by mutation type.

Variant Type	Number of Variants
Missense SNV	231
Nonframeshifting indel	5
Nonsense	4
Exonic/Splicing	4
Frameshifting indel	3
Splicing	2
Total	**249**

Table 2. Reportable incidental variants identified in 232 individuals.

Gene	Chrom-osome	Genomic Change	cDNA & Protein Change	Criteria met	Classification	Reportable?	Incidental?	Patient ancestry
LMNA	1	156106743G>A	NM_005572.3: c.1412G>A, p.Arg471His	MAF < cutoff In 3 unrelated affected individuals (Astejada et al., 2007; Cowan, Li, Gonzalez-Quintana, Morales, & Hershberger, 2010; Parks et al., 2008)	Likely pathogenic VUS	Yes	No	European
MSH2	2	47656969C>T	NM_000251.2: c.1165C>T, p.Arg389*	MAF < cutoff Protein truncation in disease caused by haploinsufficiency (S. R. Payne & Kemp, 2005)	Pathogenic	Yes	Yes	Admixed African/ European
MUTYH	1	45798269T>C	NM_001048173.1: c.583A>G, p.Ile195Val	MAF < cutoff Segregation in one affected family (Morak, Massdorf, Sykora, Kerscher, & Holinski-Feder,	Likely pathogenic VUS	No	Yes	European

				2011) In one unrelated singleton (Sieber et al., 2003)				
MYBPC3	11	47356628G>C	NM_000256.3: c.2870C>G, p.Thr957Ser	MAF < cutoff In 3 unrelated affected individuals (Ehlermann et al., 2008; Olivotto et al., 2008; Rodríguez-García et al., 2010)	Likely pathogenic VUS	Yes	No	European
MYLK	3	123512573delG	NM_053025.3: c.116delC, p.Pro39Leufs*26	MAF < cutoff Protein truncation in disease caused by haploinsufficiency (Li Wang et al., 2010)	Pathogenic	Yes	Yes	Latino

Chapter 3 Supplement

Supplementary Table 1. Variation found in 56 ACMG genes in whole exome sequences of 232 individuals.

Whole exome sequences from 232 individuals representing 89 families were analyzed for rare (MAF<0.01), nonsynonymous and splice variants and indels in the 56 genes on the ACMG list (Green et al., 2013). We identified 249 distinct variants, some of which were seen in more than one individual. The number and density of variants (number of variants/kb coding sequence) varied by gene. Some genes had better sequencing coverage than others.

Gene name	Number of variants	Variant Density (variants/kb coding length)	Average coverage (read depth)	% of UCSC-defined gene (coding exons) with > 15 reads
BRCA2	21	2.05	119.19	100
APOB	18	1.31	145.84	99.3
CACNA1S	17	3.02	77.21	100
DSP	16	1.86	139.22	99.5
RYR1	15	0.99	70.55	95.9
MYBPC3	12	3.14	73.39	96
RYR2	12	0.81	97.18	100
BRCA1	11	1.97	128.26	98.6
MYH11	9	1.52	86.16	99.3
MYH7	9	1.55	89.83	99.3
TSC2	8	1.47	68.63	99.2
DSG2	7	2.09	117.28	98.6
FBN1	7	0.81	82.62	100
MSH6	7	1.71	140.95	100
PKP2	7	2.65	61.38	85.9
MYLK	6	1.07	86.89	99.4

APC	5	0.59	138.56	100
LDLR	5	1.94	97.36	100
MUTYH	5	3.05	100.4	100
COL3A1	4	0.91	60.42	97.6
KCNH2	4	1.15	57.09	93.4
MSH2	4	1.43	69.57	96.8
TMEM43	4	3.33	69.6	100
DSC2	3	1.11	88.23	99.2
LMNA	3	1.50	65.56	93.8
SCN5A	3	0.50	94.67	100
TSC1	3	0.86	79.03	100
KCNQ1	2	0.98	51.29	90.8
MEN1	2	1.08	88.65	97.5
MYL3	2	3.40	69.42	100
PCSK9	2	0.96	54.74	97.2
RET	2	0.60	73.72	98.8
TGFBR1	2	1.32	93.89	93.5
GLA	1	0.78	80.29	100
MLH1	1	0.44	83.88	99.9
NF2	1	0.56	69.81	99.8
PMS2	1	0.39	111.39	100
RB1	1	0.36	78.21	99.5
SDHAF2	1	2.00	89.02	100
SDHB	1	1.19	66.82	100
TGFBR2	1	0.59	89.8	100
TNNT2	1	1.17	74.71	99.9
TP53	1	0.85	75.77	100
VHL	1	1.56	83.57	100
WT1	1	0.64	56.33	100
ACTA2	0	N/A	76.42	100
ACTC1	0	N/A	90.69	99.1
MYL2	0	N/A	98.89	100
PRKAG2	0	N/A	70.88	100

PTEN	0	N/A	115.28	100
SDHC	0	N/A	45.33	92.4
SDHD	0	N/A	27.01	71.8
SMAD3	0	N/A	72.32	98.9
STK11	0	N/A	60.34	97.4
TNNI3	0	N/A	72.82	100
TPM1	0	N/A	77.35	100

Supplementary Table 2. Discordant classification of variants among databases.

Three databases (HGMD (Stenson et al., 2003), ClinVar (Landrum et al., 2014), and Emory (Bean et al., 2013)) were examined for classification of the 249 distinct variants identified in 232 individuals' whole exome sequences. Of 48 variants which were classified by more than one database, 22 variants were given discordant classifications between databases. We then compared these database classifications to our variant classifications derived from the Dorschner et al. method (Dorschner et al., 2013) and found that variant classifications given by the ClinVar and Emory databases but not HGMD were generally consistent with our variant classifications. These results show the lack of consensus for variant classification methodologies among currently available databases.
(Abbreviations: DM-Disease-causing (pathological) mutation, based on information from the literature; VUS-Variant of Uncertain Significance; N/A- variant was not listed in this database).

Gene (Transcript Accession Number)	**Chrom-osome**	**Genomic Change**	**cDNA & Protein Change**	**HGMD Classification**	**ClinVar Classification**	**Emory Classification**	**Classification from Dorschner et al. Method**	**Rationale**
BRCA1 (NM_007294.3)	17	41251803T>C	c.536A>G, p.Tyr179Cys	DM	No known pathogenicity	N/A	Likely benign VUS	Seen in combination with known pathogenic mutation (Judkins et

								al., 2005)
BRCA1 (NM_007294.3)	17	41245900T>G	c.1648A>C, p.Asn550His	DM	No known pathogenicity	N/A	Likely benign VUS	Seen in combination with known pathogenic mutation (Tavtigian et al., 2006)
BRCA1 (NM_007294.3)	17	41223048A>G	c.4883T>C, p.Met1628Thr	DM	Probably not pathogenic	N/A	Likely benign VUS	Seen in combination with known pathogenic mutation (Phelan et al., 2005)
BRCA2 (NM_000059.3)	13	32907129T>C	c.1514T>C, p.Ile505Thr	DM	Probably not pathogenic	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Cavallone et al., 2010; Edwards, Kote-Jarai, Hamoudi, & Eeles, 2001)
BRCA2 (NM_000059.3)	13	32914839A>G	c.6347A>G, p.His2116Arg	DM	Probably not pathogenic	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Landrum et al., 2014)
BRCA2 (NM_000059.3)	13	32937521G>A	c.8182G>A, p.Val2728Ile	DM	Probably not pathogenic	N/A	Likely benign VUS	MAF > prevalence of disorder
BRCA2 (NM_000059.3)	13	32953971C>T	c.9038C>T, p.Thr3013Ile	DM	No known pathogenicity	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals

								(Malone et al., 2000)
BRCA2 (NM_000059.3)	13	32954034A>G	c.9101A>G, p.Gln3034Arg	DM	Variant of unknown significance	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Landrum et al., 2014)
DSC2 (NM_004949.3)	18	28672114C>T	c.304G>A, p.Glu102Lys	DM	Variant of unknown significance	N/A	Likely benign VUS	MAF > prevalence of disorder
DSG2 (NM_001943.3)	18	29099850G>A	c.166G>A, p.Val56Met	DM	Probably not pathogenic	N/A	Likely benign VUS	MAF > prevalence of disorder
DSG2 (NM_001943.3)	18	29101156T>G	c.473T>G, p.Val158Gly	DM	Probably not pathogenic	N/A	Likely benign VUS	MAF > prevalence of disorder
DSG2 (NM_001943.3)	18	29126108T>G	c.2759T>G, p.Val920Gly	DM	Probably not pathogenic	N/A	Likely benign VUS	MAF > prevalence of disorder
DSP (NM_004415.2)	6	7575014C>T	c.2422C>T, p.Arg808Cys	DM	Variant of unknown significance	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Al-Jassar et al., 2011)
DSP	6	7580795C>G	c.4372C>G,	DM	Variant of	N/A	Likely benign	MAF > prevalence of disorder

(NM_004415.2)			p.Arg1458Gly		unknown significance		VUS	
MSH2 (NM_000251.2)	2	47641430C>T	c.815C>T, p.Ala272Val	DM	Variant of unknown significance	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Landrum et al., 2014; Samowitz et al., 2001)
MYBPC3 (NM_000256.3)	11	47357479C>T	c.2686G>A, p.Val896Met	DM	Probably not pathogenic	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals (Jääskeläinen et al., 2002; Moolman-Smook, De Lange, Bruwer, Brink, & Corfield, 1999)
MYBPC3 (NM_000256.3)	11	47356628G>C	c.2870C>G, p.Thr957Ser	DM	Variant of unknown significance	N/A	Likely pathogenic VUS	MAF < prevalence of disorder In 3 unrelated affected individuals (Ehlermann et al., 2008; Olivotto et al., 2008; Rodríguez-García et al., 2010)
MYH7 (NM_000257.3)	14	23902827C>T	c.115G>A, p.Val39Met	DM	Variant of unknown	N/A	VUS	MAF < prevalence of disorder In < 3 unrelated affected individuals

					significance			(Richard et al., 2003)
MUTYH (NM_00104-8173.1)	1	45797139G>A	c.1192C>T, p.Arg398Cys	DM	Variant of unknown significance	N/A	Likely benign VUS	MAF > prevalence of disorder
PKP2 (NM_004572.3)	12	33031395G>A	c.419C>T, p.Ser140Phe	DM	Probably not pathogenic	N/A	Likely benign VUS	MAF > prevalence of disorder
RET (NM_020975.4)	10	43595999C>A	c.166C>A, p.Leu56Met	DM	No known pathogenicity	VOUS	Likely benign VUS	MAF > prevalence of disorder
RYR2 (NM_001035.2)	1	237730059C>T	c.3407C>T, p.Ala1136Val	DM	No known pathogenicity	N/A	Likely benign VUS	MAF > prevalence of disorder

Chapter 4. Mutations in *PCYT1A*, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy

Chapter 4 is reprinted from the final accepted journal article:

Hoover-Fong J, Sobreira N, Jurgens J, Modaff P, Blout C, Moser A, Kim O-H, Cho T-J, Cho SY, Kim SJ, Jin D-K, Kitoh H, Park W-Y, Ling H, Hetrick KN, Doheny KF, Valle D, Pauli RM. (2014). Mutations in *PCYT1A*, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy. *Am J Hum Genet*. 94(1): 105-12.

Julie Jurgens's Contributions: Targeted sequencing of *PCYT1A* in some families; Sanger validation and segregation analysis of variants; restriction digest validation of variant in one family; writing portions of manuscript and assistance with submission process; figure and table preparation; assistance with revision and reformatting.

Introduction

The spondylometaphyseal dysplasias (SMDs) are a group of about a dozen rare disorders characterized by short stature, irregular, flat vertebrae and metaphyseal abnormalities. Aside from spondylometaphyseal dysplasia Kozlowski type (MIM 184252) caused by mutations in *TRPV4* (MIM 605427) and spondyloenchondrodysplasia (MIM 607944) due to mutations in *ACP5* (MIM 171640), the genetic etiologies of SMDs are unknown (Warman et al., 2011). Two of these unexplained SMDs have ophthalmologic manifestations: SMD with cone-rod dystrophy (SMD-CRD [MIM 608940]) and axial SMD with retinal degeneration (MIM 602271).

Delineated clinically a decade ago, SMD-CRD is a presumed autosomal recessive disorder with postnatal growth deficiency leading to profound short stature; rhizomelia with bowing of the lower extremities; platyspondyly with anterior vertebral protrusions;

progressive metaphyseal irregularity and cupping with shortened tubular bones; and early onset, progressive visual impairment associated with a pigmentary maculopathy and electroretinographic evidence of cone-rod dysfunction (Kitoh et al., 2011; Sousa, Russell-Eggitt, Hall, Hall, & Hennekam, 2008; Turell, Morrison, & Traboulsi, 2010; Walters et al., 2004). In contrast to retinitis pigmentosa, the CRDs have early involvement of cone photoreceptors (Hamel, 2007).

Here, we report loss of function mutations in *PCYT1A* (MIM 123695) as the cause of SMD-CRD. *PCYT1A* encodes CTP:phosphocholine cytidylyltransferase (CCTα; Fagone & Jackowski, 2013; Kent, 2005), a key enzyme in the CDP-choline or Kennedy pathway for *de novo* phosphatidylcholine biosynthesis.

Subjects and Methods

We used whole exome and targeted sequencing of members of 6 unrelated families with 8 individuals with SMD-CRD (Figure 1). Three families were submitted to the Baylor-Hopkins Center for Mendelian Genomics (BHCMG) through the online submission portal PhenoDB (Hamosh et al., 2013) and, to confirm our observations in the first 3 families, we recruited 3 additional families for targeted candidate gene sequencing. The clinical features of these individuals are summarized in Table 1 and briefly reviewed here. Six of the subjects have been described in previous publications (Kitoh et al., 2011; Walters et al., 2004).

Subject 1 (BH2265_1; Family 1; Figure 2A) was reported when she was 20 y old (Walters et al., 2004). Now age 29 y, she has done well with continued linear growth to an adult height of 93.9 cm (-10.7 SD) and modest progression of limitation of range of motion. Visual impairment has not progressed since around age 10 y. Subject 2

(BH2283_1; Family 2, Figure 2E), originally described at age 11 y, is now 20 y old with a current adult height of 139 cm (-5.3 SD) and some progression of joint stiffening. His visual function declined during his second decade and he now requires low vision aids. Subject 3 (BH2233_1; Family3) is a previously unreported 61-year-old female who was first seen at age 51. Although her skeletal phenotype is similar to that of the others described here, she had a late-onset retinal phenotype (Figure 2B). Radiographs from childhood were reported to show platyspondyly and metaphyseal changes; adult radiographs show hypoplasia of the posterior vertebral bodies but no anterior vertebral protrusions. Adult height (measured at age 54 y) is 108.3 cm (-8.4 SD). Visual symptoms were not apparent until middle age. An ERG at age 43 y (performed because she had a brother with visual impairment) was said to be normal. By age 51 y, however, her ERG showed evidence of cone-rod dysfunction and her current vision is limited (Table 1). At age 54 y, her examination showed short stature, rhizomelic limb shortening, brachydactyly, stiffness of large joints and internal tibial torsion. Her family history includes an affected brother, who died at age 45 y and was known to have short stature and a confirmed CRD. Her parents may be distantly related. Subjects 4-6 were reported by Walters et al.[2] as their cases 3-5, and have not been formally assessed since that time but provided updated information regarding visual function (Table 1, Figure 2C and D). Subject 7 (Family 5) was previously described (Kitoh et al., 2011). Subject 8 (Family 6) is a 23 m old previously unreported Korean male referred for evaluation of growth failure and disproportionate shortening of the limbs. He has an increased antero-posterior thoracic diameter, rhizomelic shortening of his extremities and bilateral bowing of the legs, and mildly limited elbow extension, knee extension, and hip abduction. No visual or

hearing impairment was noted at his first examination. Linear growth was impaired (by North American standards): at 23 m he was 68.4 cm (-5.2 SD), and at 48 m 71.2 cm (-7.4 SD). From age 2 y he had frequent pneumonias with episodes of desaturation and O_2 dependency thought to be due to chest wall deformity. He had a waddling gait due to coxa vara deformity. Radiographs at 23 m showed short, bowed long bones with flared, cupped, and spurred metaphyses, and the adjacent epiphyses were large and rounded. In the hands and feet, the metaphyses of the short tubular bones had mild cupping, widening, and flaring, and the diaphyses were short. The vertebral bodies were ovoid, mildly flattened, with anterior projections. These radiographic abnormalities were more severe at age 45 m. Mild scoliosis developed. Although no visual impairment was noted at age 23 m by age 45 m fundus examination showed hypopigmented macular atrophy in both eyes with markedly decreased photopic and moderately decreased scotopic ERGs.

For our molecular studies, we isolated genomic DNA from fresh whole blood using the Gentra Puregene Kit (Qiagen Sciences, Germantown, MD). Subjects 1, 2 and 3 were genotyped on Illumina's ExomeChip1.1 GWAS array. This allows us to estimate inbreeding coefficient based on the observed and expected homozygous genotypes at genome wide level among 199 unrelated Caucasian samples (defined by PCA) after LD pruning (PLINK). The inbreeding coefficient was 0.1535, -0.007 and 0.035 for subjects 1, 2 and 3 respectively, suggesting subject 1 (BH2265_1) is the product of an unrecognized consanguineous union. Although subject 3 is also homozygous for p.A99V mutation, B allele frequency plot showed multiple loss-of-heterozygosity segments across genome for subject 1 but not subject 3. This result suggests that homozygosity for p.A99V in subject 1 is a result of consanguinity while in subject 3 it is the result of recurrent mutation (see

below).

For WES, we captured the CCDS exonic regions and flanking intronic regions totaling ~51Mb using the Agilent SureSelect Human All Exon V4 51Mb Kit and performed paired end 100 bp reads on subjects 1-3 using the Illumina HiSeq2000 platform. We aligned each read to the reference genome (NCBI human genome assembly build 37; Ensembl core database release 50_361; Hubbard et al., 2009) using the Burrows-Wheeler Alignment (BWA) tool (H. Li & Durbin, 2009) and identified single nucleotide variants (SNVs) and small insertion-deletions (indels) using SAMtools (H. Li et al., 2009). We also performed local realignment and base call quality recalibration using GATK (DePristo et al., 2011; McKenna et al., 2010). We identified potentially causal variants using standard filtering criteria: SNV and indel minimal depth of 8X, root mean square mapping quality of 25, strand bias p-value below 10^{-4}, end distance bias below 10^{-4}, and filtering out SNVs within 3bp of an indel and indels within 10bp of each other; followed by the use of the Analysis Tool of PhenoDB (Hamosh et al., 2013) to design the prioritization strategy (Sobreira, personal communication). We prioritized rare functional variants (missense, nonsense, splice site variants and indels) that were homozygous or compound heterozygous in each of the 3 subjects and excluded variants with a MAF > 0.01 in dbSNP 126, 129, and 131 or in the Exome Variant Server (release ESP6500SI-V2) or 1000 Genomes Project (1000 Genomes Project Consortium et al., 2012). We also excluded all variants found in our in-house controls (CIDRVar 51Mb). We generated a homozygous and a compound heterozygous variant list for each subject and merged them to identify genes that were mutated in both alleles in all three subjects. We designed PCR primers to amplify exons and flanking intronic splice sites followed by

direct Sanger sequencing to validate candidate causative variants, determine their segregation within families, and to sequence *PCYT1A* in subjects 4-8. All the variants described here were based on the Refgene transcript NM_005017.2 and NCBI human genome assembly build 37 (Table 1).

Results

Analysis of WES data in subjects 1-3 identified 2 genes containing candidate causal mutations in all 3 subjects: *TTN* (MIM 188840) and *PCYT1A*. *TTN* is a large gene (313 exons) expressed primarily in skeletal and cardiac muscle (Y. Zhang et al., 2007). *TTN* variants are found frequently in controls (Lopes et al., 2013) and have been implicated as causative in various cardiac and skeletal myopathies (Carmignac et al., 2007; Gerull et al., 2002; Hackman et al., 2002; D. S. Herman et al., 2012; Itoh-Satoh et al., 2002; Lange et al., 2005; Lopes et al., 2013; Satoh et al., 1999), but not in individuals with retinal or skeletal dysplasia phenotypes. For these reasons, we removed *TTN* from consideration, leaving *PCYT1A* as the only gene with rare variants in both alleles in all 3 subjects. Subjects 1 and 3 are homozygous for the *PCYT1A* missense variant p.Ala99Val (c.296C>T) in exon 5 and subject 2 is a compound heterozygote for p.Ala99Val and p.Pro150Ala (c.448C>G) in exon 6 of *PCYT1A*. These variants are not present in the >6,000 individuals in the Exome Variant Server nor in the 1092 individuals whose sequence is currently available from the 1000 Genomes. Direct Sanger sequencing of PCR amplified products confirmed appropriate Mendelian segregation of these variants in available family members (Figure 1). Based on these results, we used PCR and bidirectional Sanger sequencing to interrogate all *PCYT1A* exons and flanking intronic sequence in subjects 4-8 (3 probands and 2 affected sibs). All were compound

heterozygotes for rare *PCYT1A* variants.

Thus, in total, we identified 8 rare *PCYT1A* variants (1 nonsense, 1 frame shifting indel, and 6 missense variants) present either in the homozygous or compound heterozygous state in 8 individuals with SMD-CRD in 6 families from around the world (Table 1). The missense mutations all change highly conserved residues, including one, Ala99, that is altered by two different variants (p.Ala99Val, p.Ala99Thr) and all are predicted to be damaging by SIFT (Kumar et al., 2009) and either probably damaging or possibly damaging by PolyPhen-2 (Table 1; Adzhubei et al., 2010). The c.296C>T variant producing the p.Ala99Val change occurs at a CpG dinucleotide on the reverse strand, while, the mutation in the same codon producing p.Ala99Thr (c.295G>A) in family 4 does not involve a CpG dinucleotide. To determine if the three unrelated subjects with the p.Ala99Val change have a shared or recurrent mutation, we utilized the SNP genotyping data to analyze IBD sharing and runs of homozygosity to show that, as predicted by the different geographical origins of the 3 families, the p.Ala99Val variants in families 1-3 are on different haplotypes indicating that recurrent CpG mutations are responsible.

Discussion

Located at 3q29, *PCYT1A* (Figure 3) contains 10 exon, is ubiquitously expressed and encodes CCTα, an amphitropic enzyme that catalyzes the synthesis of CDP-choline from phosphocholine and CTP. The CCTα reaction is the rate-limiting step in the major pathway for phosphatidylcholine synthesis (Figure 4; Fagone & Jackowski, 2013; Kent, 2005). In mammals, phosphatidylcholine can also be synthesized from phosphatidylethanolamine in a reaction catalyzed by phosphatidylethanolamine N-

methyltransferase (PEMT) but the expression of the enzyme is limited to liver (Figure 4; Fagone & Jackowski, 2013; Glunde, Bhujwalla, & Ronen, 2011; Kent, 2005). Phosphatidylcholine is the predominant phospholipid in eukaryotic membranes (Fagone & Jackowski, 2013; Kent, 2005). Consistent with the importance of CCTα in phosphatidylcholine synthesis, a mouse knockout of *Pcyt1a* is an early embryonic lethal in the homozygous state (L. Wang, Magdaleno, Tabas, & Jackowski, 2005). An X-linked *PCYT1A* paralog*, PCYT1B* (MIM 604926)*,* encodes three isoforms CCTβ1, CCTβ2 and CCTβ3 with activities similar to CCTα but whose expression is limited to the central nervous system (Carter, Demizieux, Campenot, Vance, & Vance, 2008; Fagone & Jackowski, 2013).

The 367 residue CCTα has 4 domains: a 73 residue N-terminal domain containing the nuclear localizing signal, a 163 residue catalytic domain, a 64 residue membrane-binding M domain, and a C-terminal 67 residue unordered domain with multiple residues that undergo phosphorylation (Figure 3; Ding et al., 2012; J. Lee, Johnson, Ding, Paetzel, & Cornell, 2009). CCTα is found as an inactive homodimer in the nucleoplasm where, in response to lipid signaling, it binds to the nuclear membrane through interactions mediated by the amphipathic M domain and becomes catalytically active (Fagone & Jackowski, 2013; J. Lee et al., 2009). The sequence and domain structure of CCTα is highly conserved across phylogeny and the crystal structure of the catalytic domain of rat CCTα, which has 100% sequence identity with human CCTα, has been solved (J. Lee et al., 2009).

The *PCYT1A* missense alleles identified in the subjects with SMD-CRD all involve residues conserved across vertebrate and most of invertebrate and prokaryotic

phylogeny. All mutations change residues in the catalytic domain of CCTα: Ala99 (altered in subjects 1, 2, 3, 4, 5 and 6) is adjacent to Gln98 which is directly involved in binding CDP; Pro150 (altered in proband 2) is adjacent to Trp151 which is also directly involved in binding choline; Phe191 (altered in proband 8) is located 3 residues N-terminal to Thr194 which plays a role in coordinating CDP. The p.Glu129Lys and p.Arg223Ser substitutions alter residues that are linearly more distant from those that interact directly with substrates (J. Lee et al., 2009). The nonsense mutation, c.847C>T, occurs in the penultimate exon 9, 50 bp from the 3' end of the exon and possibly results in nonsense mediated RNA decay of the mutant transcript. The 1 bp indel allele in subject 7 (c.990delC) changes the frame and creates a longer transcript with 102 extra amino acids (p.Ser331Profs*140). We conclude that these variants, on the basis of the their extremely low frequency, the high conservation of the involved residues and predicted deleterious consequences for enzymatic activity, their occurrence in homozygosity or compound heterozygosity in multiple unrelated affected individuals in six families and their segregation fitting an autosomal recessive model, are *PCYT1A* loss of function alleles responsible for the SMD-CRD phenotype.

Like the deficiency of choline kinase B caused by mutations in *CHKB* (MIM 612395) responsible for muscular dystrophy, congenital megaconial type (MIM 602541; Mitsuhashi et al., 2011), SMD-CRD seems to be another inborn error in the phosphatidylcholine synthetic pathway. The central importance of this lipid in membrane biology suggests that complete loss of function of this pathway would result in embryonic lethality, as observed in mice homozygous for *Pcyt1a* null alleles (L. Wang et al., 2005). Thus, it is puzzling that individuals with SMD-CRD are healthy aside from

their skeletal and retinal involvement. One possible explanation is that despite the predicted severe functional consequences of the mutations we describe, there remains some residual CCTα function. This may be augmented by activity of the CCTβ isozymes and by synthesis of phosphatidylcholine from phosphatidylethanolamine by the PEMT catalyzed reaction in the liver (Fagone & Jackowski, 2013). The activities of the CCTβ isozymes and PEMT were increased in cells derived from conditional *Pcyt1a* knockout mice (Jacobs, Devlin, Tabas, & Vance, 2004; D. Zhang et al., 2000). These possibilities should be explored as they may suggest therapeutic strategies that could be effective, especially in younger SMD-CRD individuals.

Our results suggest a connection between bone and retinal lipid metabolism that would explain the sensitivity of these two tissues to deficiency of CCTα and impaired phosphatidylcholine synthesis. Photoreceptors have an especially high demand on membrane biosynthesis due to the daily shedding of outer segment discs (LaVail, 1976). Thus, it is not surprising that these cells might be especially susceptible to defects in the biosynthetic pathway for the most abundant membrane phospholipid. Indeed, photoreceptor degeneration in the *rd11* mouse has been shown to be due to loss of function mutations in the gene encoding lysophosphatidylcholine acyltransferase 1 (*Lpcat1*), a phospholipid remodeling enzyme (Bridges et al., 2010; Friedman et al., 2010). Interestingly, bone formation is abnormal in certain Mendelian disorders caused by mutations in genes encoding proteins involved in fatty acid metabolism, including rhizomelic chondrodysplasia punctata (MIM 215100; Braverman et al., 2002) and the Conradi-Hunermann form of chondrodysplasia punctata (MIM 302960; G. E. Herman et al., 2002). These observations suggest key functions for membrane lipids in bone

formation and homeostasis.

Among our families there was no obvious genotype-phenotype correlation apparent from evaluation of the 6 affected individuals in 4 families with mutations altering codon Ala99. Subjects 1 and 3, both homozygous for c.296C>T, p.Ala99Val, have strikingly dissimilar age at onset of visual phenotype, of clinical characteristics and of radiologic features. This broad range of phenotypic severity in individuals with the same mutations at the disease gene locus suggests that variation in other components of the phosphatidylcholine pathways and/or environmental variables may have a strong influence on phenotypic manifestations and severity. Understanding these variables may lead to improved management of individuals with SMD-CRD. Interestingly, one of the individuals described here, subject 8, has had frequent pneumonias and has required supplemental oxygen from age 2 y. We speculate that this might relate to the high demand for phosphatidylcholine in surfactant biosynthesis by the alveolar type II cells (Agassandian & Mallampalli, 2013; Goss, Hunt, & Postle, 2013). These cells are known to express *PCYT1A* at high levels and phosphatidylcholine is the major phospholipid component of surfactant (Agassandian & Mallampalli, 2013; Goss et al., 2013; Ridsdale, Tseu, Wang, & Post, 2001). A conditional, epithelial cell-specific *Pcyt1a* knockout mouse had severe respiratory failure at birth and reduced levels of surfactant (Tian, Zhou, Rehg, & Jackowski, 2007). Although only one of the individuals described here has had this problem, the potential biological connections suggest that additional studies of pulmonary function and surfactant status in proband 8 and possibly other cases of SMD-CRD is warranted.

To our knowledge, SMD-CRD is the only described disorder in which CRD is

associated with a skeletal dysplasia. By contrast, there are several examples of RP associated with bony abnormalities including some for which the responsible gene is yet to be identified (Ehara, Kim, Maisawa, Takasago, & Nishimura, 1997; Isidor, Le Merrer, Ramos, Baron, & David, 2009). CRD and RP may be part of the same phenotypic spectrum of retinal degeneration. Mutations in certain genes have been reported to cause either phenotype (Cremers et al., 1998; Yang et al., 2008). For this reason, we suggest that *PCYT1A* should be assessed in any individual with CRD or RP associated with any form of spondylometaphyseal dysplasia or spondyloepiphyseal dysplasia, including in particular, individuals with the axial SMD with retinal degeneration phenotype (MIM 602271).

Acknowledgements

We are grateful to the families for participating in this project. The Kathryn and Alan Greenberg Center for Skeletal Dysplasias and a grant from the National Human Genome Research Institute, (1U54HG006493) provided support for this work. The authors acknowledge intellectual contributions from all members of the Baylor-Hopkins Center for Mendelian Genomics (BHCMG). We also thank Dr. Michael Bamshad for making the connection of BHCMG with Dr. Park and his colleagues in Korea and Japan, Dr. Gerald A Fishman for providing fundus photographs of subject 3, and Drs. Larry Nogee and Dan Raben for helpful discussions regarding phospholipid and surfactant metabolism. We have no conflicts to disclose.

Web Resources

The URLs for data presented herein are as follows:

Baylor-Hopkins Center for Mendelian Genomics, http://mendeliangenomics.org.

Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA, http://evs.gs.washington.edu/EVS/.

1000 Genomes, http://www.1000genomes.org.

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/.

Picard software, http://picard.sourceforge.net.

PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/ibdibs.shtml.

PolyPhen-2, http://www.genetics.bwh.harvard.edu/pph2/

SIFT, http://sift.bii.a-star.edu.sg/

Chapter 4 Figures

Figure 1. Pedigrees of families 1-6 showing the segregation of *PCYT1A* mutant alleles. Alleles with the wild type genotype indicated by "+". Samples were not available for individuals lacking a genotype designation.

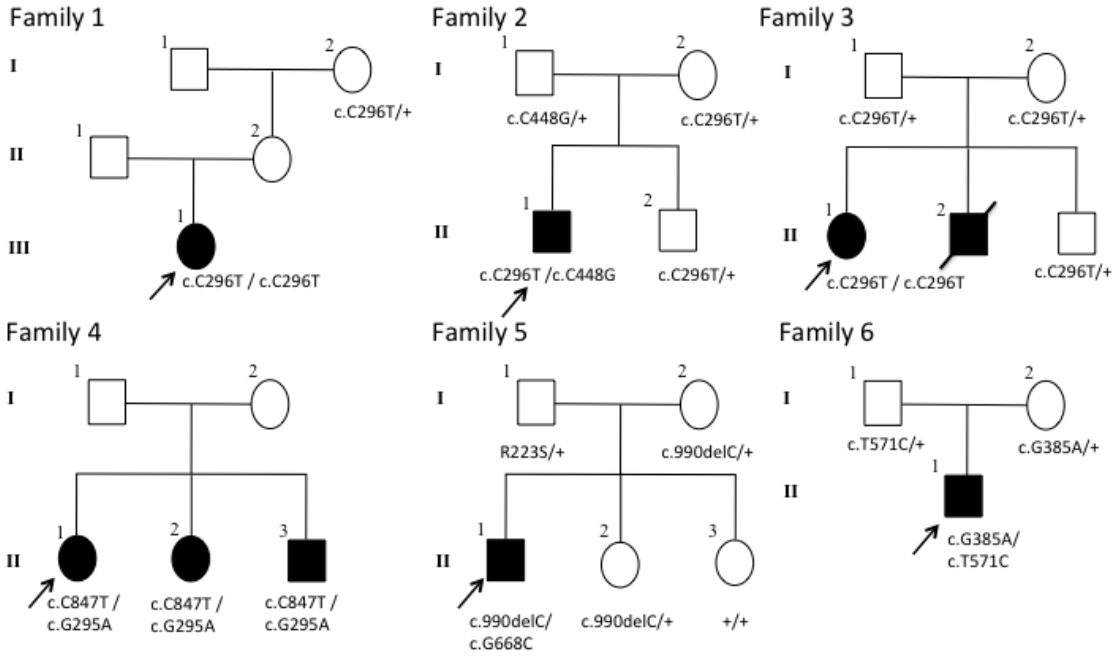


Figure 2. Clinical and radiographic features of SMD-CRD. A – General physical phenotype (Subject 1 at age 16 y 5 m); B – Fundus photograph showing pigmentary maculopathy (Subject 3 at age 61 y); C – Spine and pelvis radiograph demonstrating platyspondyly, characteristic pelvic configuration and proximal femoral metaphyseal changes (Subject 6 at 3 y 8 m of age); D – Pelvis and femurs showing the marked metaphyseal changes that are typical (Subject 4 at age 11 y 1 m); E – Knee radiographs, similarly showing the marked metaphyseal changes that are typical in this diagnosis (Subject 2 at 13 y).

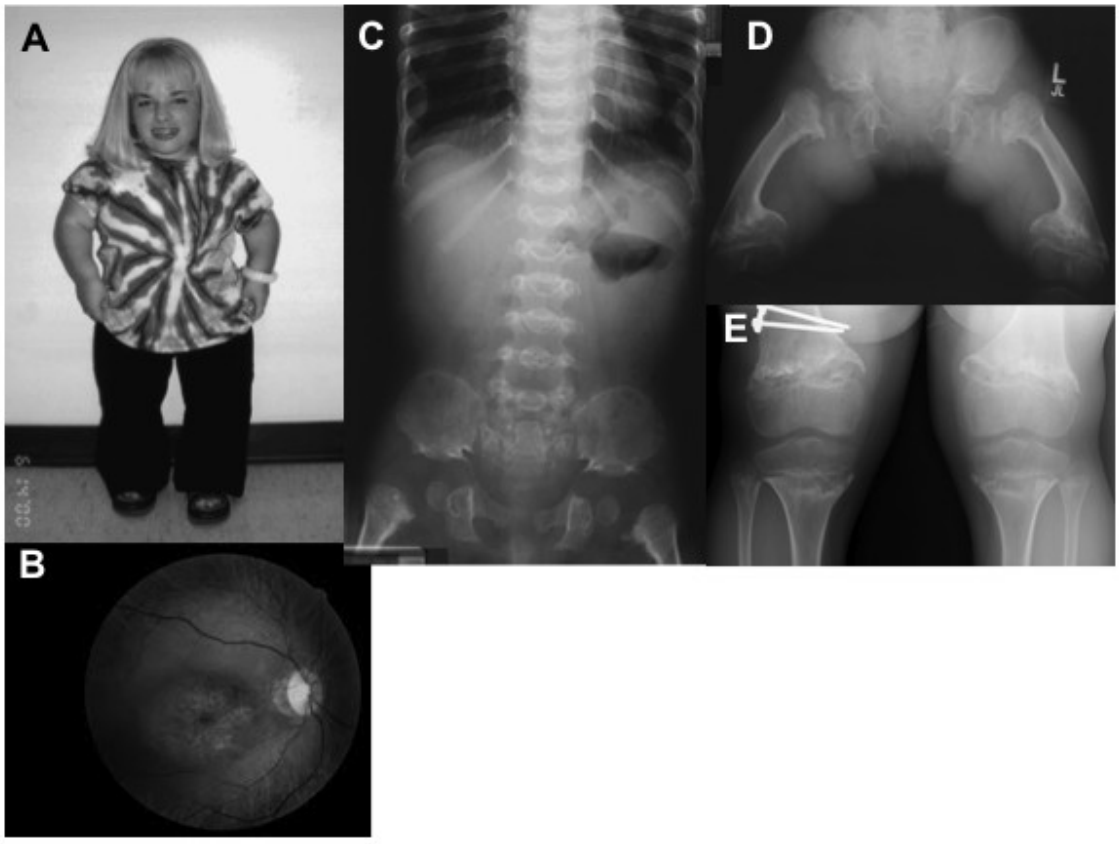


Figure 3. *PCYT1A* structure and domain organization of CCTα. *PCYT1A* is located at 3q29 and includes 10 exons (gray rectangles) as shown on the top line of the diagram. The thin rectangles encode the 5' and 3' UTR sequences. The right angle arrow denotes the start of the open reading frame and the asterisk indicates the location of the stop codon. The genomic coordinates on chromosome 3 are shown above. Below in color is a diagram of CCTα showing the domain structure with the diagonal dashed lines indicating the exons that contribute to each domain. The N-terminal domain (N) is in light blue with a dark blue rectangle indicating the position of the nuclear localization signal. The catalytic domain (C) is in pink. The amphitrophic membrane-binding domain (M) is in green and the C-terminal phosphorylated domain (P) is in purple. The numbers below indicate the residues at the boundaries of the domains. The mutations and their locations are shown below.

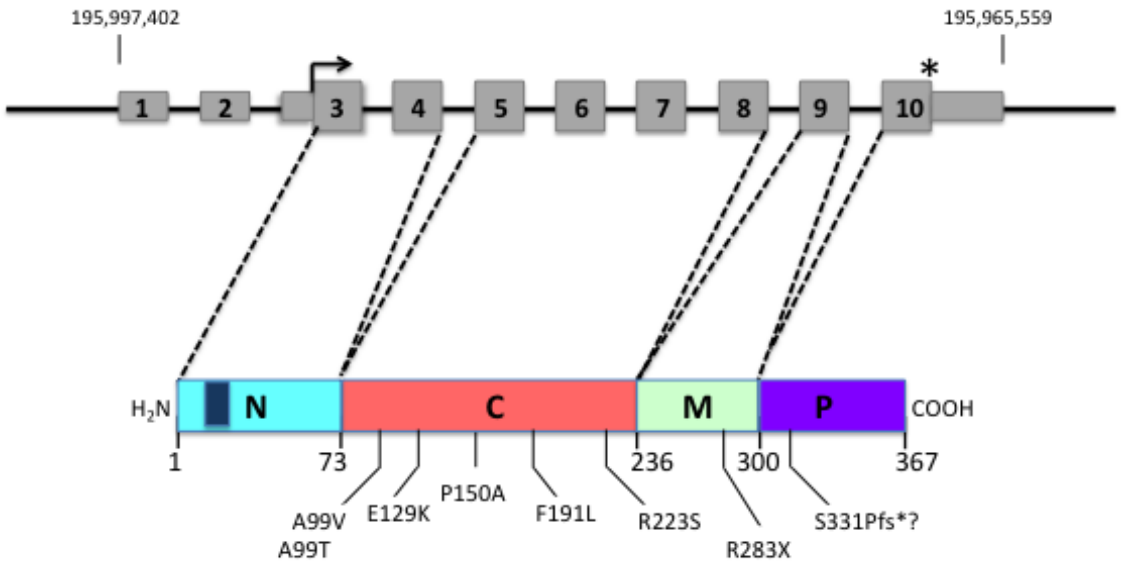
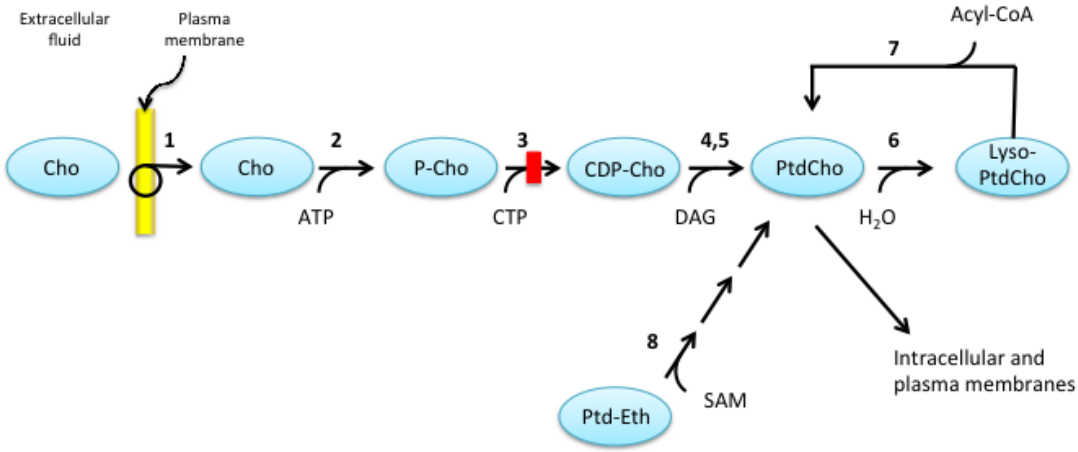


Figure 4. Pathways of phosphatidylcholine biosynthesis. In this figure, enzymes are indicated by numbers as defined below. Choline (Cho) is transported from the extracellular fluid across the plasma membrane by Na^+-dependent, ATP-requiring choline transporter-like proteins (1). Free intracellular choline is phosphorylated by choline kinase (2) to produce phosphocholine (P-Cho). The latter is converted to cytidine-diphosphate choline (CDP-Cho) in a reaction catalyzed by phosphocholine cytidylyltransferase (CCTα, 3). CDP-choline is esterified with diacylglycerol (DAG) to phosphatidylcholine (PtdCho) in reactions catalyzed either by cholinephosphotransferase (4) or choline/ethanolaminephosphotransferase (5). PtdCho is converted to lyso-phosphatidylcholine (Lyso-PtdCho) by phospholipase A (6) which, in turn, is converted back to PtdCho, in a reaction catalyzed by acylglycerophosphate acyltransferase (7). In mammals, a second pathway of PtdCho synthesis is limited to liver, where PtdCho is synthesized from phosphatidylethanolamine (Ptd-Eth) in a series of methylation reactions catalyzed by phosphatidylethanolamine *N*-methyltransferase (8) with S-adenosylmethionine (SAM) as the methyl donor. The red rectangle indicates the position of the block in PtdCho synthesis caused by deficiency of CCT.



Chapter 4 Tables

Table 1. Demographic, Clinical and Molecular Findings

	Family 1	**Family 2**	**Family 3**	**Family 4**			**Family 5**	**Family 6**
	Subject 1 (BH2265_1)	**Subject 2 (BH2283_1)**	**Subject 3 (BH2233_1)**	**Subject 4**	**Subject 5**	**Subject 6**	**Subject 7**	**Subject 8**
Previously reported	Walters et al., 2004 Case 1	Walters et al., 2004 Case 2	No	Walters et al., 2004 Case 3	Walters et al., 2004 Case 4	Walters et al., 2004 Case 5	Kitoh et al., 2011	No
Family origin	North Europe	North Europe	Greece	North Europe	North Europe	North Europe	Japan	Korea
Age of initial recognition	7 m	13 m	51 y	36 m	27 m	24 m	6 m	23 m

Best corrected visual acuity	20/100	20/180	8/80	5/250	20/250	10/250	NA	NA
Pigmentary maculopathy	+	+	+	+	+	+	+	+
Cone-rod dystrophy (age of diagnosis)	+ (13½ y)	+ (17 m)	+ (51y)	+ (36 m)	+ (27 m)	ND	+ (11 y)	+
Height SD (most recent measurement available)	-10.7 SD	-5.3 SD	-8 .4 SD	NA	-6.1 SD	NA	-7.9 SD	-7.4 SD

Bowing of long bones	+	+	+	+	+	+	+	+
Platyspondyly	+	+	+	+	+	+	+	+
Metaphyseal irregularity and cupping	+	+	+	+	+	+	+	+
PCTY1A genotype	c.296C>T/ c.296C>T	c.296C>T/ c.448C>G	c.296C>T/ c.296C>T	c.847C>T/ c.295G>A	c.847C>T/ c.295G>A	c.847C>T/ c.295G>A	c.990delC/ c.669G>C	c.385G>A/ c.571T>C
PCTY1A protein change	p.Ala99Val/ p.Ala99Val	p.Ala99Val/ p.Pro150Ala	p.Ala99Val/ p.Ala99Val	p.Arg283X/ p.Ala99Thr	p.Arg283X/ p.Ala99Thr	p.Arg283X/ p.Ala99Thr	p.Ser331Profs*140/ p.Arg223Ser	p.Glu129Lys/ p.Phe191Leu

SIFT / PolyPhen	0.04/1	0.04/1 0/1	0.04/1	0/1	0/1	0/1	0/0.791	0.02/1 0/1

Chapter 5:

Loss of function in *PCYT1A* has broad and cell-type specific consequences on lipid metabolism, chondrocyte differentiation, and lipid droplet formation

Jurgens J, Chen S, Sobreira N, Robbins S, Franca A, Dastgheyb R, Khuder S, Hoover-Fong J, Woods C, Collins F, Christodoulou J, Yamamoto G, Baratela W, Nguyen S, Haughey N, Cornell R, Valle D

Julie Jurgens's Contributions: Conceptualization and execution of all experimental procedures, data analysis and interpretation for all figures and tables, preparation of manuscript and all figures and tables

Introduction

Spondylometaphyseal dysplasia with cone-rod dystrophy (SMD-CRD) is a rare, autosomal recessive disorder of the skeleton and the retina. Clinical characteristics include progressive, early-onset photoreceptor degeneration—particularly in the macula—as well as short stature, bowing of the long bones, metaphyseal flaring, rhizomelic shortening, platyspondyly, and scoliosis. Using whole exome sequencing and Sanger sequencing, Hoover-Fong et al. (2014) and Yamamoto et al. (2014) identified homozygous or compound heterozygous variants in *PCYT1A* as the cause of SMD-CRD (Hoover-Fong et al., 2014; Yamamoto et al., 2014). These variants ranged from missense (RefSeq NM_005017.2: c.296C>T, p.Ala99Val; c.295G>A, p.Ala99Thr; c.385G>A, p.Glu129Lys; c.448C>G, p.Pro150Ala; c.571T>C, p.Phe191Leu; and c.669G>C, p.Arg223Ser) to nonsense variants (c.847C>T, p.Arg283Ter) or frameshifting indels (c.990delC, p.Ser331Profs*?; c.968dupG, p.Ser323Argfs*38; c.996delC, p.Ser333Leufs*?). Since the publication of these initial reports, our group has identified

two additional missense alleles in *PCYT1A* in patients with SMD-CRD (c.341G>C, p.Ser114Thr and c.718T>C, p.Tyr240His; Jurgens et al., personal communication).

Interestingly, compound heterozygous *PCYT1A* variants were also detected in two unrelated probands with congenital lipodystrophy and fatty liver disease (F. Payne et al., 2014). These patients had no ascertainable retinal or skeletal findings, suggesting a distinct pathophysiological mechanism. Both patients shared an in-frame deletion variant (c.838_840delCTC, p.Glu280del) in trans to a second variant in the other allele (c.424G>A, p.Val142Met or c.996delC, p.Ser333Leufs*164, respectively). Testa et al. recently reported a third class of patients with biallelic *PCYT1A* variants and retinal degeneration, but no accompanying skeletal dysplasia or lipodystrophy (Testa et al., 2017). Both probands shared the same *PCYT1A* missense variant (c.277G>A, p.Ala93Thr) in trans to an inactivating variant in the second allele (c.897+1G>A or c.847C>T, p.Arg283Ter, respectively).

PCYT1A encodes CTP:phosphocholine cytidylyltransferase α (CCTα), which catalyzes the rate-limiting step in *de novo* synthesis of phosphatidylcholine (PC) by the Kennedy pathway. *PCYT1A* is expressed ubiquitously in mice and in humans, and PC is the predominant membrane phospholipid in mammalian cells (R. B. Cornell & Ridgway, 2015; Karim, Jackson, & Jackowski, 2003; Keenan & Morré, 1970; L. Wang, Magdaleno, Tabas, & Jackowski, 2005; GTEx project release v7). Complete loss of *Pcyt1a* expression has devastating consequences: *Pcyt1a* $^{-/-}$ mice experience embryonic lethality (L. Wang et al., 2005), and MT-58 CHO cell lines with a 95% loss-of-function point mutation in *Pcyt1a* undergo apoptosis (Cui et al., 1996; Sweitzer & Kent, 1994). Based

on this evidence, PC production by the Kennedy pathway appears to be essential for life, at least in these model systems.

Though most PC is produced by the CCTα-dependent Kennedy pathway in mammalian cells, there are other compensatory pathways for PC production in mammals (R. B. Cornell & Ridgway, 2015; Jackowski et al., 2004; Fig. 2). *PCYT1B*, a paralog of *PCYT1A*, encodes the enzyme CCTβ, which catalyzes the same Kennedy pathway reaction as CCTα but has limited expression in the brains and reproductive tissues of mice and humans (Jackowski et al., 2004; Karim et al., 2003; GTEx project v7). Independently of the Kennedy pathway, PC can be formed by sequential methylation of another phospholipid, phosphatidylethanolamine (PE), by the PEMT pathway, which functions primarily in liver and adipocytes (Hörl et al., 2011; Kennedy & Weiss, 1956). Finally, in a series of reactions known as the Lands cycle, PC can either be converted into lysophosphatidylcholine (LPC) through fatty acid removal catalyzed by phospholipase A_2 (PLA_2), or LPC can be reacylated to form PC by lysophosphatidylcholine acyltransferases (LPCATs) (Lands, 1957; Vance & Vance, 2004). Interestingly, rd11 mice with *Lpcat1* deletions have a retinal degeneration phenotype, bolstering support for involvement of this arm of the phospholipid metabolism pathway in retinal degeneration (Friedman et al., 2010).

CCTα has a variety of functional domains which regulate its activity, conformation, and localization (R. B. Cornell & Ridgway, 2015). First, CCTα possesses a catalytic domain which is responsible for catalyzing PC production by the Kennedy pathway. In addition, an N-terminal nuclear localization signal localizes CCTα to the nucleus in most cells, where it exists diffusely throughout the nucleoplasm under basal conditions. Upon

activation, CCTα undergoes a series of conformational changes mediated by a membrane-binding domain (M domain) and a C-terminal phosphorylation domain in order to bind the nuclear lamina. Membrane binding elicits catalytic activation of the enzyme. It is worth noting that most SMD-CRD variants are missense variants within the catalytic domain of the enzyme (Fig. 1), although two variants fall in the M domain (p.Tyr240His and p.Arg283Ter) and three frameshifting indels localize to the C-terminal phosphorylation domain (p.Ser331Profs*?, p.Ser323Argfs*38, and p.Ser333Leufs*?).

CCTα has a number of interesting features from a cell biology standpoint, including its relationship with cellular organelles called lipid droplets (LDs). LDs form in response to loading with neutral lipids (e.g. oleate) in order to promote lipid storage and to protect against cytotoxicity (Walther, Chung, & Farese, 2017). Phospholipids such as PC are required for the formation of LD membranes, which encapsulate neutral lipids and prevent their coalescence (Krahmer et al., 2011). Some neutral lipids activate CCTα and stimulate PC production; for instance, oleate facilitates the dephosphorylation and membrane translocation of CCTα, leading to its activation (Aitchison, Arsenault, & Ridgway, 2015; Cornell & Vance, 1987; Gehrig, Cornell, & Ridgway, 2008; Lagace & Ridgway, 2005; Pelech, Pritchard, Brindley, & Vance, 1983; Y. Wang, Macdonald, & Kent, 1993). Accordingly, Drosophila, mouse, and rat cells deficient in CCTα orthologs accumulate fewer and larger lipid droplets as compared to their wild-type counterparts in response to oleate loading (Aitchison et al., 2015; Guo et al., 2008; Krahmer et al., 2011).

Here we describe the generation and characterization of three distinct cellular model systems—cultured dermal fibroblasts, HEK293 cells, and ATDC5 cells—to interrogate diverse functional consequences of SMD-CRD variants in *PCYT1A*. Using cultured skin

fibroblasts obtained from SMD-CRD patients, we sought to understand the consequences of SMD-CRD variants on cellular phenotypes in an endogenous context. Next, we generated *PCYT1A*-null HEK293 cell lines in order to assess the impact of variants on previously described cell biology phenotypes such as lipid droplet formation. Finally, we created a *PCYT1A*-null ATDC5 (pre-chondrocyte) cell lines, which are putatively more relevant to SMD-CRD pathophysiology and allowed us to measure the effects of *PCYT1A* variants on chondrocyte proliferation and differentiation. Using a combination of biochemical, cellular, and lipidomic approaches, we determined that SMD-CRD variants or complete loss of *PCYT1A* expression results in highly variable and cell-type specific phenotypic consequences.

Materials and Methods

Patient consent

Our study was approved by the Johns Hopkins Medicine Institutional Review Board and by the IRBs of other participating institutions. We obtained informed consent from all responsible individuals who participated in this study.

Cell Culture

Primary adherent fibroblast cells were derived from skin biopsies of SMD-CRD patients or wild-type controls. ATDC5 cells were a gift from Dr. Jill Fahrner of Johns Hopkins University (ECACC, Cat#99072806). HEK293T cells were a gift from Dr. Jeremy Nathans of Johns Hopkins University. CHO-K1 cells were a gift from Dr. James Phang of the NIH (ATCC, Cat#CCL61). All cell lines were cultured at 37° C under 5% CO_2 and passaged using 1 x trypsin-EDTA (Gibco, 15400-054). HEK293, CHO, and fibroblast cell lines were cultured in 1 x MEM (Gibco, 11430-030) supplemented with

non-essential amino acids (Gibco, 11140-050), 1% L-glutamine (Gibco, 25030-149), and 10% fetal bovine serum (Gemini Biosciences, 100-106). For experiments using delipidated serum, culture medium was the same, except 10% delipidized fetal bovine serum (Gemini Biosciences, Cat#900-123) was substituted for regular fetal bovine serum.

Prior to chondrocyte differentiation, ATDC5 cells were maintained in DMEM/F12 50/50 (Corning Cellgro, 15-090-CV) with 1% L-glutamine (Gibco, 25030-149), and 5% fetal bovine serum (Gemini Biosciences, 100-106), and 1% penicillin-streptomycin (Gibco, 15140-122). Chondrocyte differentiation was induced by adding 1 x insulin/transferin/selenium (Corning Cellgro, 25-800-CR) to the culture medium for 21 days.

DNA extraction and Sanger sequencing

DNA was extracted from blood or cell lines using the Puregene Blood Core Kit B (Qiagen, Cat#158467). *PCYT1A* exons were PCR-amplified from SMD-CRD patient genomic DNA with Accuprime Taq polymerase (Invitrogen, Cat#12339-016) and standard thermal cycling conditions using custom primer sequences.

***PCYT1A* cDNA construct generation**

RNA was extracted from *PCYT1A*-wild-type human fibroblasts using the RNeasy Mini Kit (Qiagen, Cat#74104), reverse-transcribed according to manufacturer protocol (SuperScriptIII One-Step RT-PCR system with Platinum Taq DNA Polymerase;Invitrogen, Cat# 12574-018), and cloned into the pcDNA3 mammalian expression vector (Invitrogen, V790-20) using primers designed to amplify full-length *PCYT1A* cDNA (based on NCBI transcript NM_005017.2) and attach restriction enzyme sites for subsequent cloning (forward primer: 5'-

TACGTAAGCTTAGCGCCACCTCAGAAGATAA-3', reverse primer: 5'-CTGAGCTCGAGTTGGGGTCACAATTTGGAAT-3'). Single base-pair SMD-CRD missense variants were then introduced using QuikChange II XL Site-Directed Mutagenesis Kit (Agilent, Cat # 200521). All plasmids were sequence-verified prior to use.

Plasmid transfection

Transfections were conducted using Lipofectamine 2000 as prescribed by manufacturer protocol (Invitrogen, Cat#11668-019) with ratios of 1.6 μg total plasmid DNA: 4 μL Lipofectamine for 12-well plate transfections, or ratios of 4.0 μg total plasmid DNA: 10 μL Lipofectamine for 6-well plate transfections. For lipid droplet rescue experiments, plasmid cDNAs harboring wild-type or SMD-CRD patient variants in *PCYT1A* were transfected into *PCYT1A*-null HEK293 cells and 24 hours post-transfection, cells were incubated with 1 mM oleate medium and further processed for imaging, as described below.

Protein extraction and western blotting

Upon reaching ~80% confluency, cells were quickly washed, scraped into ice-cold 1xPBS supplemented with protease inhibitors (Sigma-Aldrich, P8340-5ML), and centrifuged for 5 minutes at 1,000 x g. Cells were then resuspended in RIPA buffer (150 mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) with protease inhibitors (Sigma-Aldrich, P8340-5ML), vortexed briefly, and lysed at 4° C with rotation and periodic low-speed vortexing. Cell debris were pelleted via centrifugation at 12,000 x g for 15 minutes, and the supernatant was isolated for subsequent analysis. We measured total protein concentrations of cell lysates with the

BCA protein assay kit (Pierce, 23225) and diluted 40 μg of each protein with XT sample buffer (Bio-Rad, 1610791) and XT reducing agent (Bio-Rad, 1610792) as prescribed by company protocols. Samples were denatured by boiling for 5 minutes, then cooled and run on 4-12% Bis-Tris Criterion XT precast gels (Bio-Rad, 3450123) with XT MOPS Running Buffer (Bio-Rad, 1610788). Transfer was conducted onto Immun-Blot PVDF membranes (Bio-Rad, 162-0177) at 4° C with 100 V for 1 hour in 1 x Tris/Glycine Buffer (Bio-Rad, 1610734) and 20% methanol (Fisher Scientific, A41220). Blots were blocked for 1 hour at room temperature in 5% nonfat dried milk in TBST (0.5% Tween-20, 137 mM NaCl, 200 mM Tris, pH 7.5), then incubated in the appropriate primary antibodies (anti-β-actin antibody AC-15, ThermoFisher Scientific AM4302, 1:10,000 dilution; anti-CCTα, Abcam, ab109263, 1:1,000 dilution) diluted in 5% nonfat dried milk in TBST at 4° C overnight. Membranes were washed 3 x 10 minutes in TBST, incubated for 1 hour at room temperature in the appropriate secondary antibodies diluted 1:10,000 in 5% nonfat dried milk in TBST (goat anti-rabbit IgG-HRP, Santa Cruz Biotechnology sc-2004; goat anti-mouse IgG-HRP, Santa Cruz Biotechnology sc-2005). Membranes were washed 3 x 10 minutes in TBST, then incubated in ECL reagent (GE Healthcare Life Sciences, RPN2106) for 5 minutes and exposed to CL-X Posure Film (Thermo Scientific, 34091).

Phosphatidylcholine incorporation assay

To assess the effect of *PCYT1A* variants on the rate of phosphatidylcholine synthesis by the Kennedy pathway, we monitored continuous incorporation of [Methyl-^{3}H]-choline chloride (NEN Radiochemicals, NET109001MC) into phosphatidylcholine over time as described previously (Kitos, Drobnies, Ng, Wen, & Cornell, 2006). Wild-type or SMD-CRD patient fibroblast cell lines were cultured in 1xMEM with 10% FBS, 1% L-

glutamine, and non-essential amino acids in 60 mm dishes until ~80% confluency, and then counted using the Beckman Coulter Particle Counter Z1. Normal growth medium was aspirated, and cells were incubated in serum-free medium at 37° C with 5% CO_2 for 1 hour. We then added 5 μCi [^{3}H]-methyl-choline per dish and returned to the incubator, or quenched immediately for our 0 minute time point. After each time point was reached, cells were quickly rinsed in ice-cold 1xPBS, and radiolabel uptake was quenched by adding 1 mL of ice-cold methanol per dish. Cells were then placed on ice and scraped into prechilled glass tubes containing 1 mL chloroform, and residual cells were scraped with an additional 1 mL of methanol. We then added 0.75 mL H_20 per tube, mixed by vortexing at low speed, and added 1 mL additional chloroform followed by 1 mL of H_20 per tube. Cells were then vortexed briefly to mix, followed by centrifugation for 5 minutes at 1000 RPM to form a biphasic mixture. The lower organic phase (containing lipids) was then transferred to a scintillation vial, evaporated under nitrogen gas (Organomation Associates, Inc., Berlin, MA, USA) and resuspended in 10 mL Insta-Gel Plus scintillation fluid (Perkin Elmer, 6013399), followed by counting with a Beckman LS6500 scintillation counter. Previous thin layer chromatography analyses have shown that >90% of counts in the organic phase following up to 2 hours incubation are attributable to phosphatidylcholine (M. N. P. Ng, Kitos, & Cornell, 2004), so counts in this phase were used as a proxy for radiolabeled phosphatidylcholine levels. All sample counts were corrected for background counts, obtained using a blank scintillation vial.

Genome editing

Genome editing was performed as described previously, with slight modification (Moyer & Holland, 2015). Guide RNAs against target genomic sequences from human or

mouse *PCYT1A* (RefSeq NC_000003.11 or NC_000082.5, respectively) were designed using the CRISPR MIT portal (Hsu et al., 2013), modified into oligonucleotides for cloning purposes, and cloned according to the submitter's laboratory protocol into the pSpCas9(BB)-2A-Puro (PX459) V2.0 plasmid (Addgene, Cat #62988), which was a gift from Feng Zhang's laboratory (Ran et al., 2013). By combining two guide RNAs targeting the beginning and end of the genomic sequence of either human or mouse *PCYT1A*, we generated ~31 kb genomic deletions in either human (HEK293 cells) or mouse (ATDC5 cells) *PCYT1A*, effectively excising the majority of the coding sequence. Cells were seeded for transfection at ~2 x 10^5 cells/well in six-well plates, grown to ~80% confluency, and transfected using Lipofectamine 2000 as prescribed by manufacturer protocol (Invitrogen, 11668-019), with equal DNA amounts of each guide RNA-encoding plasmid (2 μg each plasmid DNA: 10 μl Lipofectamine). Next, we treated cells with puromycin-containing medium at either 2 μg/mL (HEK293 cells) or 4 μg/mL (ATDC5 cells) 24 hours post-transfection. Medium was replaced daily for 2-3 days, until cells in untransfected control wells died. Single cells were then isolated by serial dilution in puromycin-free medium, expanded into isogenic clones, subjected to DNA extraction, and screened for both wild-type and deletion-containing genomic DNA by PCR amplification. Clones which contained the *PCYT1A* deletion but not wild-type genomic DNA were genotyped using Sanger sequencing and tested for CCTα protein levels by western blot analysis.

Cell proliferation assays

Cell proliferation was measured using the Cell Counting Kit-8 (Dojindo Molecular Technologies, CK04-05), as prescribed by manufacturer protocol. Cells were counted

using the Beckman Coulter Particle Counter Z1, plated at 1000 cells per well in 96-well plates, and cultured for the respective amounts of time indicated in our growth curve. Growth was measured by incubating with CCK-8 solution for 3 hours, followed by absorbance reading at 450 nm with the Biotek Synergy 2 plate reader and Gen5 software v2.01.14.

Oleate incubation for lipid droplet and nuclear lamina translocation studies

In order to induce lipid droplet formation, oleic acid: BSA complexes were prepared for addition to cell culture medium. First, we prepared a 100 mM sodium oleate solution (Sigma-Aldrich, cat#O3880-1G) in distilled water and sonicated on ice (Branson Sonifier 250 with tapered microtip, Fisher Scientific, Cat#22-309782) under a constant duty cycle and output level of 2 for approximately 2 minutes, or until solution reached clarity. We then mixed this solution with a 200 mg/ml fatty acid-free BSA (Sigma-Aldrich, A8806-5G) in PBS and briefly sonicated again until solution was clear. This generated a final stock of 100 mg/mL BSA with 10 mM oleate, which we diluted 10-fold in serum-free medium (1 x MEM supplemented with non-essential amino acids and 1% L-glutamine) for a 1 mM working oleate solution. Cells were incubated in 1 mM oleate medium for various timepoints prior to processing for immunofluorescence or lipid droplet staining.

Cell staining and immunofluorescence

Cell culture was performed using the conditions described above in 12-well plates containing sterile 18-mm glass cover slips (VWR, 48380-046). Cells were fixed for 10 minutes in 4% paraformaldehyde (Sigma-Aldrich, P6148-500G), permeabilized with 0.2% Triton X-100 (Sigma-Aldrich, T-8532) in 1 x PBS for 10 min, and blocked in 1% BSA

(Sigma-Aldrich, A-9647) in 1 x PBS for 1 hour, all at room temperature. Cells were then processed for immunofluoresence or lipid droplet staining.

For immunofluorescence experiments, cells were incubated in primary antibody (anti-CCTα, Abcam, ab109263) diluted 1:200 in 1% BSA in PBS overnight at 4°C. Cells were then incubated in Alexa Fluor 488 or 555 secondary antibodies (Thermo Fisher Scientific, A-21428 or A-11008) diluted 1:300 in 1% BSA in PBS for one hour at room temperature. For lipid droplet staining, cells were incubated for one hour in BODIPY 493/503 dye (ThermoFisher Scientific, Cat#D-3922, stored as 1 mg/mL stock in 100% ethanol at -20°C) diluted in 1:500 in 1 x PBS.

Cell nuclei were counterstained with DAPI (Life Technologies, D1306) diluted to 300 nM in 1 x PBS for 10 minutes at room temperature, followed by mounting with ProLong Gold Antifade Mountant (Thermo Fisher Scientific, P36930) on glass slides (Fisher Scientific, 12-550-15). Images were obtained using the Zeiss LSM 510 Meta Confocal Microscope under 63x oil magnification.

Lipid Droplet quantification

Lipid droplet sizes and numbers were quantified as described previously (Aitchison et al., 2015). Using Image J software (v1.47), we converted images to 8-bit, adjusted image threshold, and used the "analyze particle" command with a particle size distribution of 20 to infinity pixels^2 set to exclude on edges.

Alcian blue staining

Alcian blue was performed as previously described (Lu et al., 2016), with slight modification. Upon reaching ~80% confluency, cells were rinsed in 1xPBS, fixed with 4% paraformaldehyde for 10 minutes at room temperature, and then incubated in 1% Alcian

Blue 8GX (Sigma-Aldrich, A5268-10G) pH 1.0 at 4° C overnight. After briefly washing in 1xPBS, cells were lysed in 1% SDS for one hour and absorbance was measured at 605 nm using the Biotek Synergy 2 plate reader with Gen5 software v2.01.14.

Lipid standards for mass spectrometry

The following lipid standards were obtained from Avanti Polar Lipids (Alabaster, AL): 1,2-dilauroyl-*sn*-glycero-3-phosphate (sodium salt) (PA 12:0/12:0, PA C12), 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (PC 12:0/12:0, PC C12), 1,2-dilauroyl-*sn*-glycero-3-phosphoethanolamine (PE 12:0/12:0, PE C12), 1,2-dilauroyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (sodium salt) (PG 12:0/12:0, PG C12), 1,2-dilauroyl-*sn*-glycero-3-phospho-L-serine (sodium salt) (PS 12:0/12:0, PS C12), cholesteryl-d7 palmitate (Cholesterol-d7 ester 16:0, Cholesterol ester d7), N-lauroyl-D-*erythro*-sphingosine (C12 Ceramide d18:1/12:0, Cer C12), N-heptadecanoyl-D-erythro-sphingosine (C17 Ceramide d18:1/17:0, Cer C17), N-lauroyl-D-erythro-sphingosylphosphorylcholine (12:0 SM d18:1/12:0 SM C12), 1,3-dihexadecanoyl glycerol (d5) [1,3-16:0 DG (d5), DG d5], 1,3(d5)-dihexadecanoyl-2-octadecanoyl-glycerol [TG d5-(16:0/18:0/16:0), TG d5], D-galactosyl-ß-1,1' N-lauroyl-D-*erythro*-sphingosine [C12 Galactosyl(ß) Ceramide (d18:1/12:0), GlcCer C12], D-lactosyl-ß-1,1' N-dodecanoyl-D-*erythro*-sphingosine [Lactosyl (ß) C12 Ceramide, LacCer C12]. We purchased APCI positive calibration solution from AB Sciex (Cat. #4460131; Concord, Ontario, Canada).

Stock solutions of PC C12 (12.5 mg/mL), PS C12 (5 mg/mL), and SM C12 (5 mg/mL) were dissolved in methanol. Stock solutions of Cer C12 (2.5 mg/mL), Cer C17 (2.5 mg/mL), PE C12 (6.25 mg/mL), PA C12 (3.125 mg/mL), PG C12 (6.25 mg/mL), DG d5

(0.25 mg/mL), TG d5 (0.25 mg/mL), LacCer C12 (0.625 mg/mL) and GlcCer C12 (1.25 mg/mL) were dissolved in dichloromethane/methanol 1:1 (v/v). Finally, cholesterol ester d7 was dissolved in dichloromethane. All solutions were stored at -20°C until use. Ultrapure water (resistivity > 18 MΩ cm) was used for mass spectrometry experiments.

Cell culture and lipid extraction for mass spectrometry

For mass spectrometry experiments, we measured various lipid classes in *PCYT1A* wild-type versus *PCYT1A*-null HEK293 and ATDC5 cells, or in cultured dermal fibroblasts obtained from *PCYT1A* wild-type controls or SMD-CRD patients homozygous for the p.Ser323Argfs*38 variant in *PCYT1A*. All cells were cultured at 37° C under 5% CO_2 and passaged using 1 x trypsin-EDTA (Gibco, Cat. #15400-054). HEK293 and fibroblast cell lines were cultured in 1 x MEM (Gibco, Cat. #11430-030) supplemented with non-essential amino acids (Gibco, Cat. #11140-050), 1% L-glutamine (Gibco, Cat. #25030-149), and 10% fetal bovine serum (Gemini Biosciences, Cat. #100-106). ATDC5 cells were maintained in DMEM/F12 50/50 (Corning Cellgro, Cat. #15-090-CV) with 1% L-glutamine (Gibco, Cat. #25030-149), and 5% fetal bovine serum (Gemini Biosciences, Cat. #100-106), and 1% penicillin-streptomycin (Gibco, Cat. #15140-122). Upon reaching ~80% confluency, cells were scraped in ultrapure water, pelleted at 1000 x g, resuspended in ultrapure water, and subjected to total protein concentration measurements using the BCA protein assay kit (Cat. #23225, Pierce). Cell extracts were stored at -80°C until further processing.

We isolated lipids from cell extracts using a modified Bligh-Dyer procedure. 200 µg protein from each cell suspension was transferred into 1.5 mL plastic centrifuge tubes and adjusted to 200 µL total volume with ddH_2O. We then sonicated cell solutions to form

homogenates by applying short bursts with a sonic dismembrator (Fisher Scientific, Waltham, MA). Homogenates were transferred to glass tubes and gently mixed with 800 µL of ddH_2O, for a final volume of 1 mL. We next formed a monolayer by adding 2.9 mL methanol/dichloromethane (2:0.9, v/v) containing internal standards for 12 lipid classes to each sample. We then added 1 mL of ddH_2O and 0.9 mL dichloromethane to obtain a biphasic mixture, which was incubated at 4^{o}C for 30 min and centrifuged at 4 ^{o}C for 10 min at 3000 x g to separate organic and aqueous phases. 1 mL of the organic phase was transferred to a 2 mL glass vial and stored at -20^{o}C until use. We finally dried 0.5 mL of the organic layer extract using a nitrogen evaporator (Organomation Associates, Inc., Berlin, MA, USA) and re-suspended in 135 µL of running solvent (dichloromethane:methanol (1:1) with 5 mM ammonium acetate) containing 5 mg/mL of ceramide (C17:0) as an internal standard. All solvents used were HPLC grade.

Untargeted lipid analysis by MS/MS^{ALL}

Mass spectrometry analyses were performed via MS/MS^{ALL} on a TripleTOF^{TM} 5600 mass spectrometer (AB SCIEX, Redwood City, CA). We applied 50 µL of each lipid extraction to the system by direct infusion using a DuoSpray electrospray ionization source and autosampler (Shimazu, Canby, OR) at a flow rate of 7 µL/min. Samples were run in duplicate in the positive ion mode. A dichloromethane: methanol solution (1:1, v/v) with 5 mM ammonium acetate was used as the running solvent. We used a mass resolution of ~30 000 for TOF MS scans and ~15 000 for product ion scans in high sensitivity mode, and automatically calibrated using APCI positive calibration solution delivered via calibration delivery system (AB Sciex) after every 10 sample injections. We used the following source parameters: ion source gases 15 psi (GSI), 20 psi (GS2),

curtain gas 30 psi, temperature 150^{o}C, positive ion spray voltage +5200 V, declustering potential of 80V and collision energy of 10V. Initial TOF MS scanning provided an overview of total lipid content at an accumulation time of 5s. Precursor ions were selected by sequential 1 Da mass steps from 200.050 to 1200.050 m/z. Analytes in each 1 Da step were introduced into the collision chamber, and fragments were identified by TOF with a scan range of 100-1500 m/z (accumulation time of 300 ms). The collision energy for each MS/MS step was 40eV. In addition, a pooled sample containing extracts from all groups was run 8 times for lipid identification and subsequent selection of targeted lipids. All data was acquired using Analyst 1.7 TF (AB SCIEX, Concord, ON, Canada).

Mass spectrometry data processing and analysis

The TOF MS and MS/MS^{ALL} data obtained from individual samples and pooled samples were aligned post-acquisition to internal standards using Analyst 1.7 TF. We conducted broad assignment and identification of lipids on the aligned data from pooled samples using LipidView^{TM} software (v1.3 Beta, AB SCIEX, Concord, ON, Canada). To reduce false-positive identifications due to instrumental noise, lipids whose MS/MS fragment peaks appeared in at least 7 of 8 pooled runs were considered for further analysis. We then calculated the coefficient of variation (CV) of the fragment peak intensity for each lipid relative to the internal standard for this class of lipid obtained over multiple runs. Lipids with CVs <20% were considered reliable and used as targeted lipids in order to identify lipids of these classes in experimental samples using LipidView^{TM}. We calculated CV values and generated targeted lipid lists using a custom in-house Matlab program. MultiQuant software (version 3.0, AB SCIEX, Concord, ON, Canada)

was used to identify lipids in experimental samples by comparing to the targeted lipid list described above. Peak intensities were corrected using internal standards for each class of lipid, and we averaged duplicates within each sample for further analysis. To ensure that data could be statistically analyzed, any zero-value intensities were replaced by 0.000001. We then proceeded with additional data analysis using R 3.4.2 (Bioconductor). Data were log-transformed, z-score normalized, and exponentially transformed so that all values were positive. In order to compare each lipid between *PCYT1A* wild-type and variant cell lines of each group (ATDC5, HEK293, or fibroblast cell lines), we used Welch's two-sample *t*-test.

Statistical procedures

All statistical analyses were performed using R 3.2.4 (Bioconductor). Multiple group comparisons were done using Tukey's Honest Significant Difference test. No outliers were removed for statistical analyses.

Results

Measurement of CCTα steady state levels and PC incorporation rates

Western blot analysis of CCTα steady state levels in cultured fibroblasts of five SMD-CRD patients (4 probands and one affected sibling) revealed reduced CCTα steady state levels in all patient cell lines as compared to wild-type controls, with the level of reduction varying according to patient genotype (Figure 3A). Fibroblasts homozygous for the p.Ala99Val variant had the mildest reductions in CCTα steady state levels (~75% wild-type levels), followed by p.Glu129Lys homozygotes (~30% wild-type levels). Fibroblasts homozygous for a 1-bp frameshifting insertion (p.Ser323Argfs*38) had only

~10% of wild-type CCTα levels, and similar levels were observed in cells compound heterozygous for the p.Ser323Argfs*38/p.Ser114Thr variants.

To assess PC synthesis, we measured continuous incorporation of [^{3}H]-choline into PC in intact fibroblasts over 2 hours. Consistent with the results of our western blotting experiments, p.Ala99Val homozygotes had the mildest impairments in PC incorporation (54% of wild-type), whereas homozygotes for the p.Ser323Argfs*38 or p.Glu129Lys alleles had only 22% of wild-type incorporation levels, and compound heterozygotes for the p.Ser323Argfs*38/p.Ser114Thr alleles had 33% of wild-type incorporation.

CCTα membrane translocation in putative cell models

In order to assess the utility of several cell models for studying CCTα variants, we first analyzed CCTα localization in response to oleate treatment. Previous studies have shown that oleate stimulation can induce CCTα translocation to the nuclear lamina in CHO cells and in HEK293 cells, although results are variable depending on the report and the cell line tested (Aitchison et al., 2015; Gehrig et al., 2008; Lagace & Ridgway, 2005). Under our experimental conditions, oleate failed to stimulate translocation of CCTα to membranes in wild-type fibroblasts or in CHO cells, but did induce membrane translocation in wild-type HEK293 cells. Based on this finding, we decided to move forward with additional cellular modeling in HEK293 cells to examine consequences of *PCYT1A* variants on cellular phenotypes.

Lipid droplet analyses

In cells from Drosophila, mice, and rats, deficiency of CCTα orthologs results in formation of fewer and larger lipid droplets than in control cells in response to oleate exposure (Aitchison et al., 2015; Guo et al., 2008; Krahmer et al., 2011). Based on this,

we hypothesized that SMD-CRD patient fibroblast cells would also develop fewer and larger lipid droplets in response to oleate treatment. However, this was not the case. The sizes and numbers of lipid droplets in both wild-type and SMD-CRD patient fibroblast cell lines were highly heterogeneous, and SMD-CRD patient fibroblasts (p.Ser323Argfs*38 homozygotes) failed to show significant differences in lipid droplet numbers or sizes as compared to control cells (Fig. 4A and 4C). Interestingly, however, *PCYT1A*-null HEK293 cells generated fewer and larger lipid droplets than their wild-type counterparts in response to oleate loading (Fig. 4B and 4D). Transfection with wild-type, p.Ala99Val, and p.Tyr240His *PCYT1A* cDNAs, but not empty vector cDNA, rescued lipid droplet phenotypes in *PCYT1A*-null cells (Fig. 4E).

Proliferation studies in HEK293 cells

Interestingly, *PCYT1A*-null HEK293 cells proliferated at normal rates as compared to wild-type cells (Fig. 5A). We hypothesized that this might be attributable to exogenous serum lipid uptake in *PCYT1A*-null cells, so we next determined the effects of culturing cells in delipidized serum. Both wild-type and *PCYT1A*-null HEK293 cells grew slower in delipidized serum as compared to normal serum, but surprisingly *PCYT1A*-null HEK293 cells actually grew more effectively than their wild-type counterparts in delipidized serum (Fig. 5A).

Effects of *Pcyt1a* knockout on chondrocyte proliferation and differentiation

To evaluate the utility of ATDC5 pre-chondrocytes as a model system, we first assessed their steady-state CCTα levels at baseline and over the course of 21 days of chondrocyte differentiation by western blot analysis. CCTα protein was expressed robustly in ATDC5 cells prior to chondrocyte differentiation and maintained stable

expression over the 21-day differentiation timecourse (Fig. 6A). Next, we generated *Pcyt1a*-null ATDC5 cell lines using genome editing and assessed the consequences of *Pcyt1a* loss-of-function on chondrocyte proliferation and differentiation. Similar to our observations in HEK293 cells, *Pcyt1a*-null ATDC5 cells proliferated at normal rates as compared to their wild-type counterparts (Fig. 5B). However, *Pcyt1a*-null ATDC5 cells actually differentiated into chondrocytes more quickly than their wild-type counterparts (Fig. 6B).

Untargeted lipidomic studies

To assess global changes in lipid metabolism in SMD-CRD patient-derived skin fibroblasts (p.Ser323Argfs*38 homozygotes) and in *PCYT1A*-null HEK293 or ATDC5 cells as compared to their wild-type counterparts, we conducted untargeted mass spectrometry on ~1000 distinct lipids. In total, the levels of ~75-200 lipids were significantly increased or decreased in the *PCYT1A*-null or SMD-CRD condition for each cell line tested, and the specific lipids altered were largely variable depending on the cell type tested, with one exception. The levels of several individual LPCs (Fig. 7A) were significantly decreased in *PCYT1A*-null HEK293 cells (LPCs 16:1, 18:1, 18:2, 20:0, 20:3, 20:4, 20:5, 22:5, and 22:6) and in SMD-CRD patient fibroblasts (LPCs 16:0, 16:1, 16:3, 18:1, 18:2, 20:1, 20:4, 22:4, and 22:5). Notably, levels of five identical LPCs (LPCs 16:1, 18:1, 18:2, 20:4, and 22:5) were reduced in both *PCYT1A*-null HEK293 cells and in SMD-CRD patient fibroblast cell lines as compared to controls.

We first examined whether the proportions of specific lipid classes in each cell line was altered in *PCYT1A*-null or SMD-CRD conditions as compared to wild-type controls. We grouped individual lipids in each cell line into the following categories: cholesterol

ester, ceramide, desmosteryl ester, diacylglycerol, galactosylceramide, lactosylceramide, lysophosphatidylcholine, lysophosphatidylethanolamine, lysophosphatidylglycerol, lysophosphatidylserine, monoakyldiacylglyceride, phosphatidic acid, phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, phosphatidylserine, sphingomyelin, or triacylglyceride. Some lipid classes were not represented in all cell lines tested. In *PCYT1A*-null ATDC5 cells as compared to their wild-type counterparts, we observed significant decreases in the overall proportions of lysophosphatidylcholines (4-fold) and lysophosphatidylethanolamines (2-fold). In SMD-CRD patient cells homozygous for the p.Ser323Argfs*38 variant as compared to wild-type controls, we observed significant increases in the overall proportions of diacylglycerol (1.9-fold) and phosphatidylethanolamine (1.6-fold). Finally, in *PCYT1A*-null HEK293 cells as compared to their wild-type counterparts, we observed significant decreases in the proportions of lysophosphatidylcholine (2-fold) and phosphatidylcholine (2.5-fold), as well as significantly increased proportions of galactosylceramide (1.9-fold), lactosylceramide (2-fold), and lysophosphatidylserine (2.5-fold).

Next, we analyzed whether the ratios of specific lipid classes were altered in *PCYT1A*-null or SMD-CRD patient cells as compared to wild-type controls for each cell type. First, we found that sphingomyelin:ceramide (SM:Cer) ratios were not significantly altered in any of the *PCYT1A*-null or SMD-CRD patient cell lines as compared to their wild-type counterparts (data not shown). However, phosphatidylcholine:phosphatidylethanolamine (PC:PE) ratios were reduced significantly by 2-3 fold in *PCYT1A*-null HEK293 cells and in SMD-CRD patient fibroblast cell lines as compared to their wild-type counterparts (Fig. 7B).

Discussion

Through western blot analyses PC incorporation assays, we determined that all SMD-CRD patient-derived dermal fibroblasts tested were hypomorphic to varying degrees, ranging from 10-75% of wild-type steady state CCTα protein levels and 22-54% of wild-type PC incorporation. Surprisingly, the level of reduction in steady-state CCTα or PC incorporation does not appear to correlate with the severity of patients' phenotypes. Fibroblasts homozygous for p.Ala99Val had the mildest biochemical phenotype of all variants tested, both in their CCTα steady state levels (75% of wild-type) and PC incorporation (54% of wild-type). Because the p.Ala99Val allele decreased PC incorporation more severely than steady-state expression, we conclude that this allele reduces the catalytic activity of CCTα independently of its effects on protein production or stability. This conclusion is further supported by the location of the affected amino acid in the crystal structure of the orthologous rat CCTα protein: the alanine residue falls adjacent to a glutamine which is involved in CDP coordination, and hence is important for catalytic activity (J. Lee et al., 2009).

Other alleles reduced steady state levels of CCTα and PC incorporation much more severely: for instance, fibroblasts homozygous for the p.Ser323Argfs*38 allele had only 10% of the steady state CCTα and 22% of the PC incorporation levels of wild-type fibroblasts. We did not expect this allele to result in complete loss of CCTα expression, since the generated termination codon is C-terminal to the penultimate exon junction and therefore is not predicted to be subject to nonsense-mediated mRNA decay. However, the strong decrease in steady state CCTα coupled with less severe reductions in PC incorporation suggest that this allele results in an unstable protein with normal or perhaps

increased catalytic activity resulting from its modified C-terminus, which typically regulates activation of the enzyme.

Compound heterozygous fibroblasts for p.Ser323Argfs*38/p.Ser114Thr alleles had only 14% of wild-type steady state CCTα but 33% of wild-type PC incorporation. The protein levels in this cell line were only slightly higher than those of p.Ser323Argfs*38 homozygotes, indicating that the p.Ser114Thr allele has similarly adverse consequences on protein production or stability, at least when in combination with the p.Ser323Argfs*38 allele. However, similar to p.Ser323Argfs*38 homozygotes, the residual CCTα protein in compound heterozygotes seems to have normal or even increased activity as compared to wild-type protein. Finally, fibroblasts homozygous for p.Glu129Lys had ~30% of wild-type steady-state CCTα levels and similar levels of PC incorporation (22% of wild-type), indicating that this variant protein is unstable and has normal or only mildly reduced CCTα activity. This glutamic acid residue falls in an α-helix at the CCTα dimerization interface which is part of a larger series of α-helices and β-strands comprising a Rossman fold domain (J. Lee et al., 2009). It is possible that this variant destabilizes CCTα by decreasing its capacity to dimerize.

Contrary to our expectations based on prior studies in Drosophila, mouse, and rat cell lines (Aitchison et al., 2015; Guo et al., 2008; Krahmer et al., 2011), SMD-CRD patient fibroblasts (p.Ser323Argfs*38 homozygotes) failed to show significant differences in lipid droplet numbers or sizes as compared to controls, indicating that oleate loading does not induce an abnormal lipid droplet phenotype in cells with this genotype. Of note, p.Ser323Argfs*38 homozygous fibroblasts had the most severe reductions of all SMD-CRD patient cells tested in both steady-state levels of CCTα (~10% wild-type levels) and

in phosphatidylcholine incorporation (~22% wild-type levels), so the absence of a lipid droplet phenotype in our experiments was surprising. This may be because the mechanism for regulating lipid droplet sizes and numbers is not conserved in human fibroblasts. Alternatively, it is possible that complete ablation of CCTα protein expression is required for lipid droplet phenotypes to manifest; indeed, SMD-CRD patient fibroblasts retained some residual CCTα expression and activity, which could be sufficient to normalize lipid droplet formation. Consistent with previous publications, *PCYT1A*-null HEK293 cells generated fewer and larger lipid droplets than did wild-type HEK293 cells in response to oleate loading, suggesting again that CCTα-mediated lipid droplet regulation may be cell-type specific in cultured human cell lines, or that complete ablation of *PCYT1A* expression may be required in order for lipid droplet phenotypes to manifest. Transfection with wild-type, p.Ala99Val, and p.Tyr240His *PCYT1A* cDNAs, but not empty vector cDNA, rescued lipid droplet phenotypes in *PCYT1A*-null cells. This may indicate that these variants in the catalytic or membrane-binding domains of CCTα provide ample enzyme activity and PC production to generate lipid droplets normally; this interpretation would be consistent with our findings in SMD-CRD patient fibroblasts, in which even small residual amounts of functional CCTα are sufficient to regulate lipid droplet formation normally. However, we also recognize that cDNA expression was driven by a CMV promoter in our transfection studies, which may simply normalize lipid droplet phenotypes by *PCYT1A* overexpression.

Given that *Pcyt1a*-null mice are embryonic lethal and CHO MT-58 cells undergo apoptosis (Cui et al., 1996; Sweitzer & Kent, 1994; L. Wang et al., 2005), we were intrigued that *PCYT1A*-null HEK293 and ATDC5 cells were grossly morphologically

normal and proliferated at normal rates as compared to their wild-type counterparts. However, these results are consistent with findings that *Pcyt1a*-null mouse peritoneal macrophages grow normally under standard cell culture conditions due to compensatory upregulation of a specific isoform of *Pcyt1b* (D. Zhang et al., 2000). Interestingly, however, these *Pcyt1a*-null mouse macrophages had increased susceptibility to cell death upon loading with free cholesterol, suggesting that additional stressors may be required to evoke cell death in CCTα-deficient cell lines. Furthermore, we found that *PCYT1A*-null HEK293 cells proliferated more quickly than their wild-type counterparts when cultured in delipidized serum. This suggests that growth rates are normalized independently of exogenous lipid uptake from serum in these cells, and that *PCYT1A*-null HEK293 cells may have a growth advantage over their wild-type counterparts when grown in conditions devoid of exogenous lipids, perhaps by upregulation of compensatory pathways for PC production.

Using untargeted mass spectrometry, we were able to investigate changes in lipid species apart from those directly implicated in the Kennedy pathway. We found that a broad array of lipids are increased or decreased in *PCYT1A*-null HEK293 or ATDC5 cells or SMD-CRD patient fibroblast cell lines as compared to controls, and the altered lipids were often cell-type specific. However, we noticed a few trends that replicated in more than one of the three cell lines tested. First, total PC:PE ratios were significantly decreased in *PCYT1A*-null HEK293 and in SMD-CRD patient fibroblast cell lines as compared to controls. Decreased PC:PE ratios have previously been associated with the loss of membrane integrity and decreases in membrane potential in certain cell types (Z.

Li et al., 2006), so it is possible that decreased ratios of these lipids have broad adverse consequences.

In addition, we found that several LPCs were significantly decreased in both *PCYT1A*-null HEK293 and in SMD-CRD patient fibroblast cell lines as compared to the corresponding control cell lines. Interestingly, five of these decreased LPCs had the same acyl chain compositions in both cell lines: LPC 16:1, 18:1, 18:2, 20:4, and 22:5. Decreased LPCs could be attributable either to increased conversion of LPC into PC by LPCAT or to decreased conversion of PC into LPC by PLA_2. Previous studies in COS cells and in CHO cells have shown that CCTα overexpression increases LPC levels by promoting PLA_2-mediated PC catabolism (Walkey, Kalmar, & Cornell, 1994; Barbour, Kapur, & Deal, 1999), so it is possible that decreased activity by this pathway is also responsible for decreases in LPC in our cell lines. However, recent studies have also demonstrated increased *Lpcat4* mRNA expression and activity toward LPC acyl chains 18:1, 18:2, 20:4, and 22:6 during the late stages of chondrocyte differentiation in ATDC5 cells (Tabe et al., 2017). Notably, LPCs with three of these four acyl chains (18:1, 18:2, and 20:4) were decreased in SMD-CRD patient fibroblast cell lines, and all four were decreased in *PCYT1A*-null HEK293 cells as compared to wild-type controls in our study. The mean levels of these and other individual LPCs were also decreased in *Pcyt1a*-null as compared to wild-type ATDC5 cells, but none of these decreases approached statistical significance. This may be due in part to the small sample size of ATDC5 cells measured in our study, so we suggest further investigation of LPCs in this cell type for future studies. Observed decreases in individual LPC species were also consistent with

decreases in overall LPC content in both *PCYT1A*-null HEK293 cells and ATDC5 cells as compared to wild-type controls.

Decreased LPCAT expression or activity may also be supported by our finding that *Pcyt1a*-null ATDC5 cells have increased rates of chondrocyte differentiation as compared to wild-type controls. *Lpcat4* knockdown has been associated with decreased rates of chondrocyte differentiation in ATDC5 cells (Tabe et al., 2017), so it is possible that compensatory upregulation of *Lpcat4* expression or activity contributes to the increased rate of chondrocyte differentiation in *Pcyt1a*-null ATDC5 cells. Changes in chondrocyte differentiation rates have also been observed in previous studies of other skeletal dysplasias. For instance, ATDC5 cells expressing the variant responsible for achondroplasia have slowed rates of chondrocyte differentiation, whereas mice null for parathyroid hormone-related protein have accelerated rates of chondrocyte differentiation (Lanske et al., 1996; Matsushita et al., 2013). Our results suggest that SMD-CRD may be another phenotype for which perturbed rates of chondrocyte differentiation contribute to skeletal dysplasia.

The overall proportions of lipids represented in each SMD-CRD variant or *PCYT1A*-null cell line were also altered, but the nature of the alterations varied among cell types. As mentioned previously, the only lipid whose overall cellular content was decreased in two cell types was LPC, which was present in lower proportions in *PCYT1A*-null HEK293 and ATDC5 cell lines as compared to corresponding wild-type controls. In *Pcyt1a*-null ATDC5 cells, we also observed decreases in overall LPE content, suggesting potential upregulation of LPE acyltransferase (LPEAT) or downregulation of PLA_2. In *PCYT1A*-null HEK293 cells, we observed decreased PC content, suggesting that

compensatory pathways for PC production may not be as robust in this cell type as in others. *PCYT1A*-null HEK293 cells also had increased content of LPS, galactosylceramide, and lactosylceramide. However, the reasons or implications for increased levels of these lipids are unclear. Both galactosylceramide and lactosylceramide are formed from ceramide, which can combine with PC to form sphingomyelin in a reaction catalyzed by sphingomyelin synthase. Perhaps decreased PC production by the Kennedy pathway promotes downregulation of sphingomyelin synthase in an attempt to maintain cellular PC, and in turn promotes increases in other ceramide metabolic pathways such as those required for galactosylceramide and lactosylceramide production. In homozygous p.Ser323Argfs*38 fibroblasts, the proportions of DAG and PE were increased as compared to wild-type fibroblasts. Increased proportions of DAG might be expected, given that SMD-CRD cells have decreased production of CDP-choline, which normally combines with DAG to form PC in the final step of the Kennedy pathway. However, increased proportions of PE were not expected *a priori* and suggest that in SMD-CRD fibroblasts, upregulation of the PEMT pathway is not likely a major means for compensatory PC production.

Overall, our results indicate that partial or complete loss of *PCYT1A* in ATDC5, HEK293, or fibroblast cell lines leads to decreased PC synthesis by the Kennedy pathway and cell-type specific secondary dysregulation of diverse additional lipid metabolic pathways, which manifests in lipid droplet phenotypes in HEK293 cells but not SMD-CRD patient fibroblasts. Our results also suggest involvement of *Pcyt1a* in chondrocyte differentiation but not proliferation or morphology. We suggest that a careful side-by-side assessment of all *PCYT1A* alleles implicated in SMD-CRD, isolated retinopathy, and

lipodystrophy in a range of cell types would be informative for dissecting further tissue-specific consequences of these alleles.

Acknowledgements

We are grateful to our patients and their families for participation in this study. We would like to thank Dr. Jill Fahrner of Johns Hopkins University for providing her ATDC5 cells, and Drs. Dan Raben and Kathy Wilson of Johns Hopkins University for helpful insights into lipid metabolism and nuclear biology, respectively. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The GTEx data used for the analyses described in this manuscript were obtained from the V7 release of the GTEx portal (https://www.gtexportal.org/), dbGaP Accession phs000424.v7.p2, on 12/04/17. Our work was supported in part by grants from the US NIH/NHGRI (T32GM07814; 1U54HG006542) and NIH/NEI (5T32EY007143-20).

Chapter 5 Figures

Figure 1. Diagram of all known SMD-CRD alleles to-date. Both previously described and novel SMD-CRD patient variants are mapped to their corresponding CCTα protein domains. Note that most variants are missense and fall within the catalytic domain. NLS-nuclear localization signal; M-membrane-binding domain; P-phosphorylation domain.

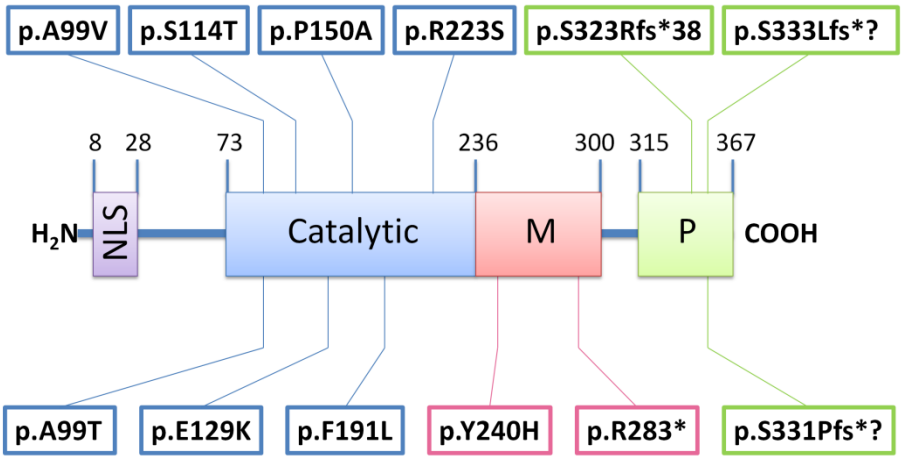


Figure 2. Kennedy pathway and related pathways for PC production. Choline (Cho) is taken in exogenously from the diet and transported across the plasma membrane into the cytosol, where it becomes phosphorylated by choline kinase to produce phosphocholine (P-Cho). Phosphocholine is converted into CDP-choline (CDP-Cho) in a reaction catalyzed by CTP:phosphocholine cytidylyltransferase (CCT), which is combined with diacylglycerol (DAG) to form phosphatidylcholine (PC). Phospholipase-A_2 (PLA_2) converts PC into lyso-phosphatidylcholine (LPC), which can be converted into PC once more in a reaction catalyzed by lysophosphatidylcholine acyltransferase (LPCAT). PC can also be derived from sequential methylation of phosphatidylethanolamine (PE) by phosphatidylethanolamine *N*-methyltransferase. The red rectangle indicates the location of the block in PC synthesis in SMD-CRD. Phosphatidylethanolamine can also be interconverted to lysophosphatidylethanolamine by phospholipase-A_2 (PLA_2) and LPE acyltransferase (LPEAT) enzymes.

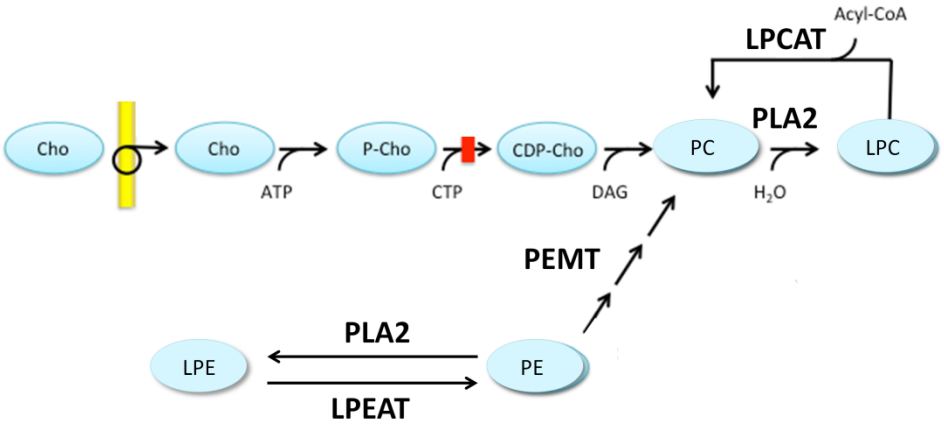


Figure 3. Western blot and PC incorporation assays of SMD-CRD patient fibroblasts. A) Representative immunoblot showing reduced protein expression of CCTα relative to loading control β-actin in cultured skin fibroblasts of five SMD-CRD patients as compared to wild-type controls, with the level of reduction varying according to patient genotype. Protein levels ranged from ~10-75% of wild-type CCTα. B) To assess PC synthesis, we measured continuous incorporation of [^{3}H]-choline into PC in intact fibroblasts over 2 hours. Cell lines again had variable PC incorporation levels, ranging from 22-54% of wild-type. Error bars represent standard deviation. n=3 biological replicates. Tukey's Honest Significance Test. *p<0.05.

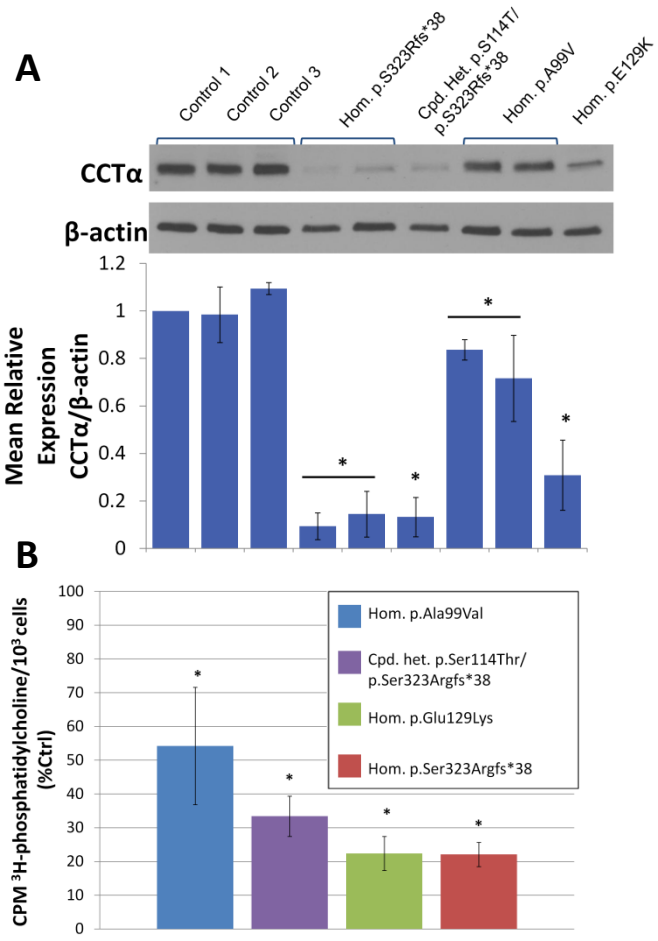


Figure 4. Lipid droplet experiments. Lipid droplet (LD) formation was induced by oleic acid loading and visualized with BODIPY staining. LD size and number were quantified using "analyze particle" command in ImageJ (v1.47). A and B) Neither LD size (p=0.43) nor number (p=0.87) differed between control and SMD-CRD patient fibroblasts. C and D) *PCYT1A*-null HEK293 cells have fewer and larger LDs than control cells. E) LD sizes and numbers are rescued by transfection with *PCYT1A* WT, p.Ala99Val or p.Tyr240His cDNAs but not corresponding empty vector controls. n=200 cells from 4 experimental trials. Error bars represent SD. Welch Two Sample t-test for panels A and C, and Tukey's Honest Significant Difference test for panel E. *p<0.05.

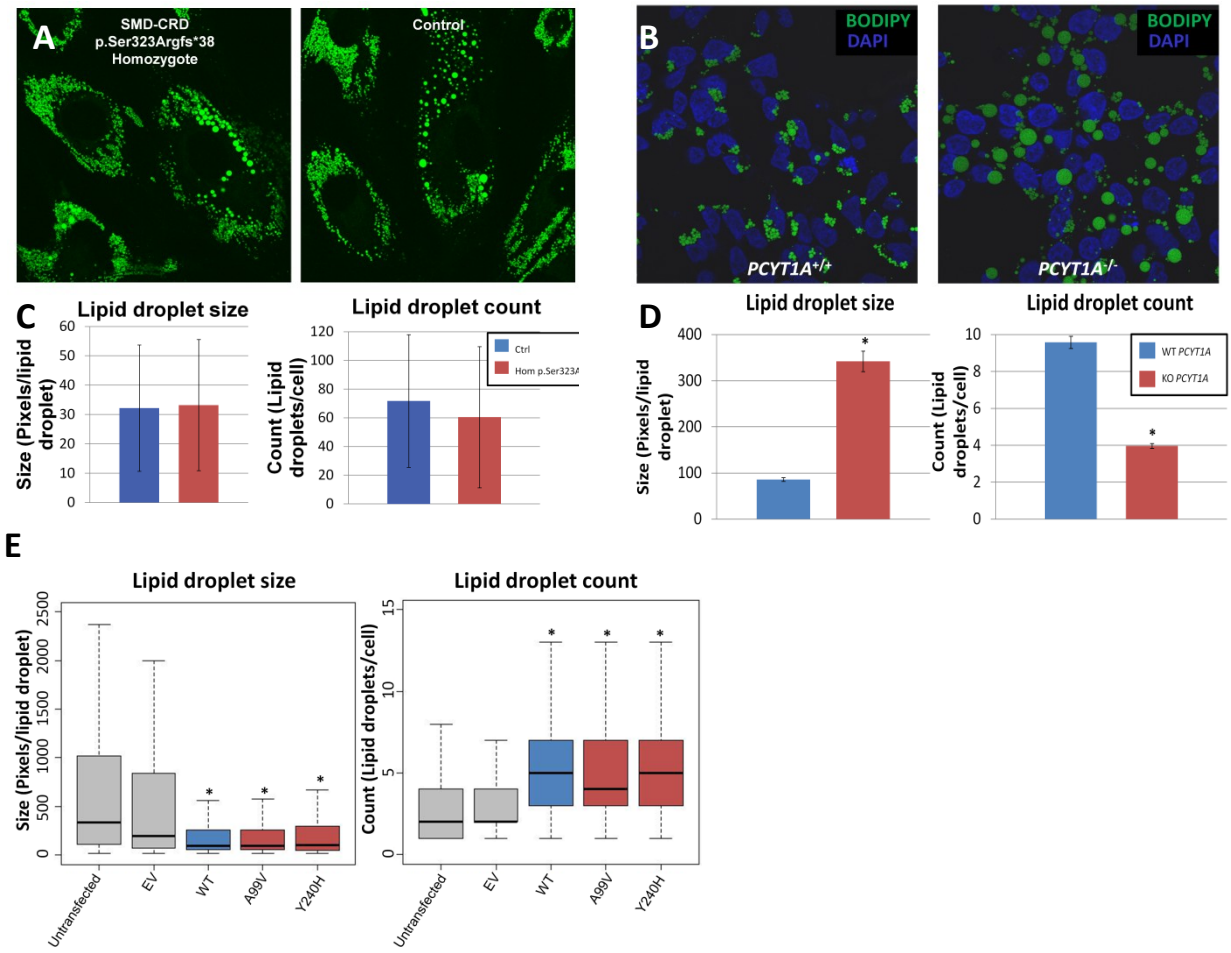


Figure 5. *PCYT1A*-null HEK293 and ATDC5 cells have normal proliferation rates. A) Wild-type and *PCYT1A*-null HEK293 cells were grown in either normal serum or delipidized serum up to 144 hours and counted every 24 hours. Growth rates were comparable for both cell lines under normal serum; however, *PCYT1A*-null HEK293 cells actually grew more rapidly in delipidized serum than their wild-type counterparts. B) Undifferentiated wild-type or *Pcyt1a*-null ATDC5 cells were grown for 5 days and counted every 24 hours using absorbance-based cell counting. At all time points, wild-type and *Pcyt1a*-null cells had comparable rates of growth.

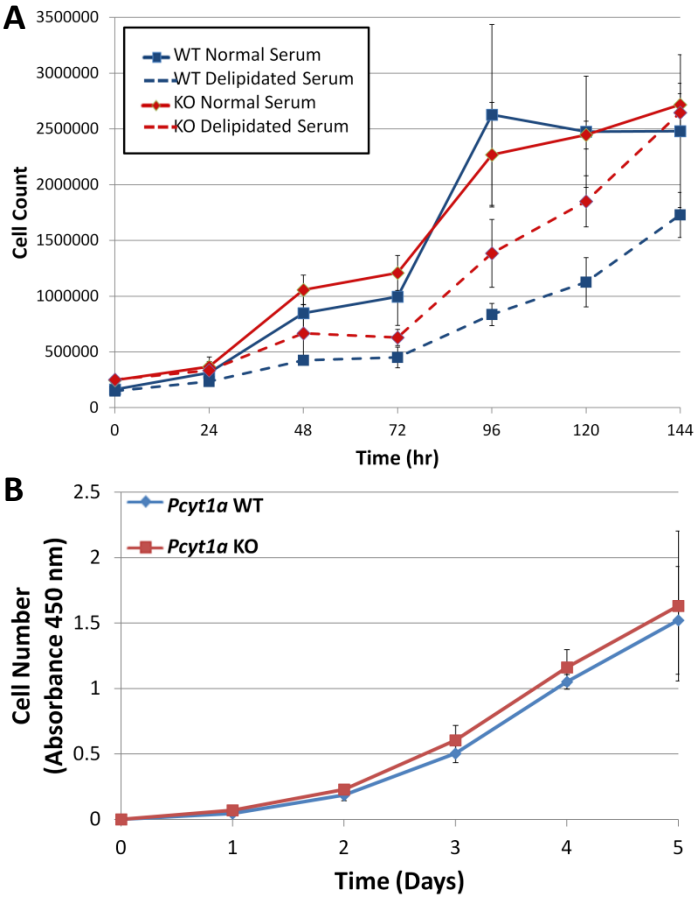


Figure 6. Involvement of Pcyt1a in ATDC5 chondrocyte differentiation. A) Western blot analysis shows that CCTα steady-state levels remain stable over the course of 21 days of chondrocyte differentiation. B) Rates of chondrocyte differentiation are increased in *Pcyt1a*-null ATDC5 cells, as measured by absorbance reading of Alcian blue staining. n=3 biological replicates for each time point.

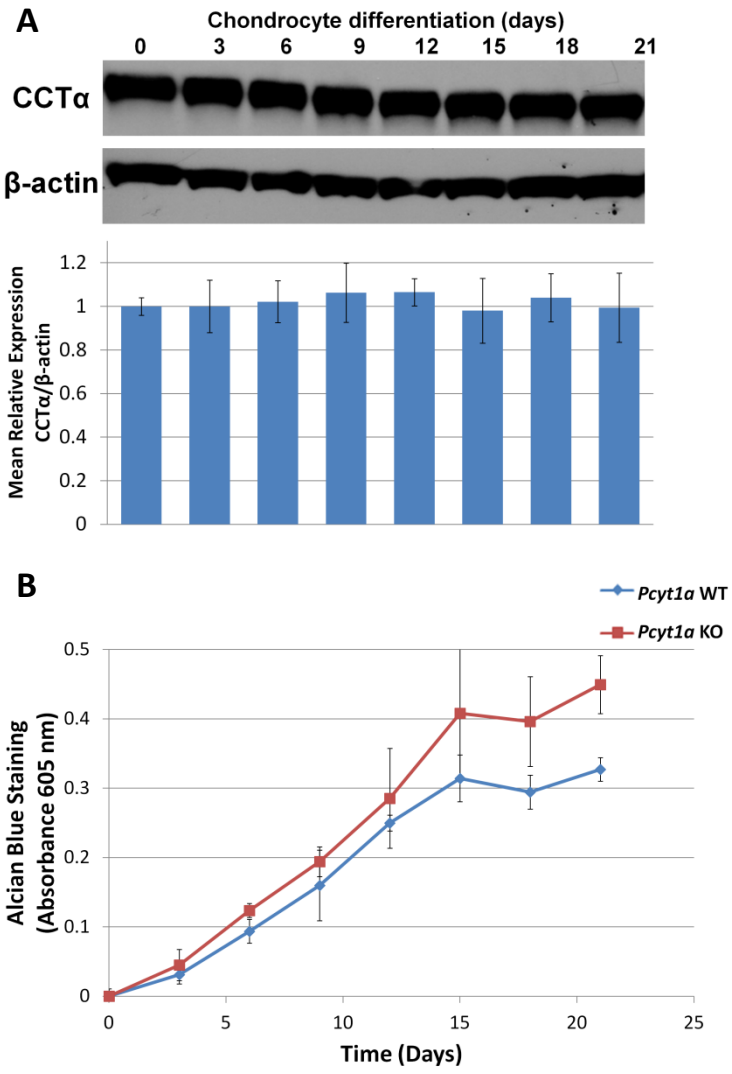
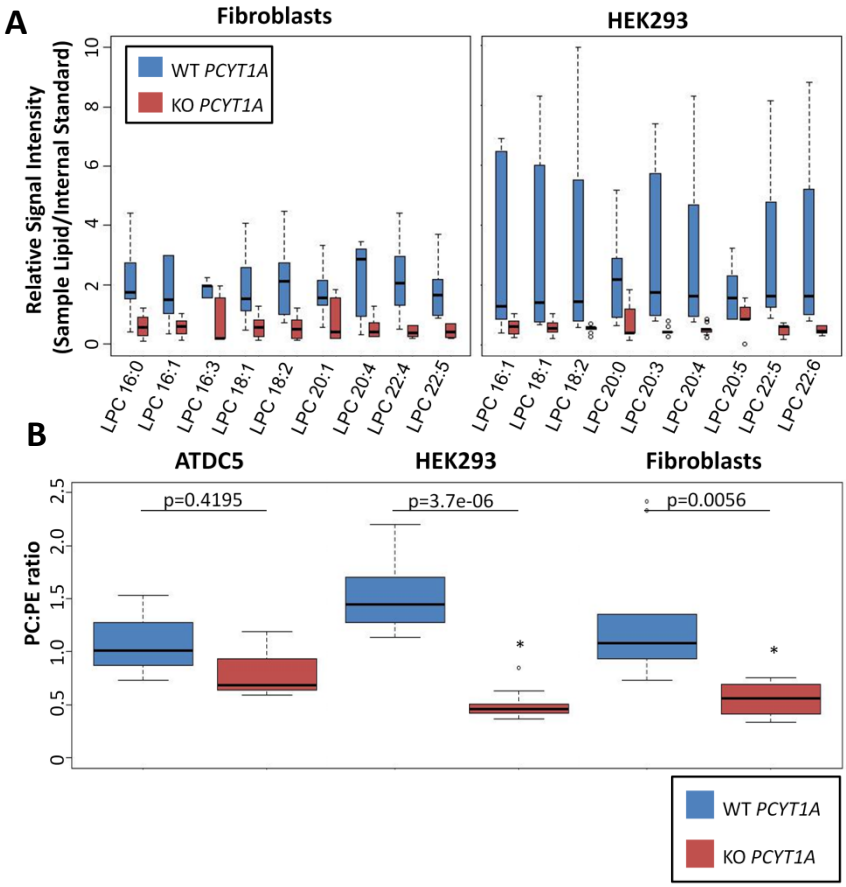


Figure 7. Untargeted lipidomics reveal decreases in LPCs and in PC:PE ratio in SMD-CRD patient fibroblasts and *PCYT1A*-null HEK293 cells. A) Several specific lysophosphatidylcholine (LPC) species (acyl chain compositions denoted) were decreased in both SMD-CRD patient fibroblasts and *PCYT1A*-null HEK293 cells as compared to wild-type controls. B) Ratios of total phosphatidylcholine (PC) to phosphatidylethanolamine (PE) were decreased in both SMD-CRD patient fibroblasts and *PCYT1A*-null HEK293 cells but not *Pcyt1a*-null ATDC5 cells as compared to wild-type controls.



Chapter 6: Conclusion

From this work, we have seen that numerous challenges exist for the detection of variants from next-generation sequencing studies as well as the functional characterization studies that follow in their wake. However, as time has progressed, novel methodological developments and analytic strategies have become available which increase our ability to identify and characterize genetic variants underlying Mendelian disorders. For instance, long-read sequencing methodologies are evolving and becoming increasingly more affordable, so we can sequence larger blocks of DNA and better detect structural variants, copy number variants, or disease haplotypes. Whole genome sequencing is also becoming more cost-effective, so as we approach saturation for variant detection in the exome, we can turn toward genome sequencing for families whose causal variants were not detectable via exome sequencing. In addition, methods for variant detection are progressing—particularly for copy number and structural variants—allowing us to revisit existing sequence data with new tools. Alternatively, we can couple existing methods such as linkage analysis with next-generation sequencing in order to solve a greater proportion of cases.

Researchers and clinicians are increasingly recognizing that families with each rare Mendelian disorder are limited, as are the resources with which to study them. Consequently, collaborative efforts are increasing, both through traditional mechanisms (see Chapter 2 for an example) and through tools such as GeneMatcher which facilitate genetic matchmaking on a larger scale (Sobreira, Schiettecatte, Valle, & Hamosh, 2015). Moreover, many groups are now integrating mRNA expression data from individual RNA-seq experiments or publicly available databases (e.g. Allen Brain Atlas and the

GTEx project) as well as proteomic studies to prioritize genes which are mutated and also have altered expression profiles (Genotype-Tissue Expression Project, www.gtexportal.org; Allen Brain Atlas, www.brain-map.org).

Large databases harboring variant frequency information (e.g. ExAC and gnomAD) as well as variant-associated phenotypic information (e.g. ClinVar) have enhanced our ability to filter variants based on their population frequencies or known clinical associations (Landrum et al., 2014; Lek et al., 2016). Many countries whose populations were previously underrepresented in genetic databases are also undertaking large-scale sequencing projects, which will enrich our understanding of human variation more globally. Moreover, our knowledge of which genes are essential for life in humans or other model organisms is also evolving through resources like the knockout mouse project and the human knockout project (Lloyd, 2011; Saleheen et al., 2017), which demonstrate the consequences of deleting specific genes. The Precision Medicine Initiative is also underway, which will elucidate relationships between genomics and phenotypic outcomes as they evolve over time (Terry, 2015).

Moreover, new computational and statistical methods are being developed that can be applied to sequencing studies. Some of these tools include improved splice site detection algorithms (M. Lee et al., 2017), catalogs of noncoding genetic elements such as those developed by the ENCODE project (Dunham et al., 2012), and tools for the spatial clustering of de novo missense variants (Lelieveld et al., 2017). In addition, we can apply more robust statistical methods such as genetic burden tests for cohort analyses (S. Lee, Abecasis, Boehnke, & Lin, 2014). Although statistical tools for detecting

oligogenic contributors to Mendelian disease have not yet been developed, such an advance is anticipated going forward.

Various methods are currently available for prioritization of candidate genes and variants, such as genic intolerance and CADD scores (Kircher et al., 2014; Petrovski et al., 2013), and additional scores have been developed for applications such as ranking variants on a genome-wide level or based on the protein domains in which they fall (Gussow, Petrovski, Wang, Allen, & Goldstein, 2016; Smedley et al., 2016; Wiel, Venselaar, Veltman, Vriend, & Gilissen, 2017).

As a more robust catalog of candidate genes and variants is developed, we will be faced with the need to characterize their functional consequences at the bench. As demonstrated in Chapter 5, this can be quite challenging and may reveal unanticipated mechanistic complexities that take time to unravel. However, several tools now exist that improve our ability to functionally characterize variants and may ease this process in the future. For instance, we can now generate knockout or knock-in models fairly quickly using CRISPR/Cas9-mediated genome editing. This method may ameliorate the challenges of allelic heterogeneity by enabling us to model specific alleles, or even to examine the functional consequences of all possible changes in a gene through saturation genome editing (Findlay, Boyle, Hause, Klein, & Shendure, 2014). Genome editing can be applied to various model systems, including cell lines, animal models, and induced pluripotent stem cells, which together give us a broadly expanded repertoire for understanding the functional consequences of genetic variation.

In the end, next-generation sequencing endeavors, increasingly sophisticated analytic strategies and functional follow-up studies will have far-reaching consequences

for patients with Mendelian disorders, their families, and society at large. We now have a wide array of disorders which we can target with pharmaceutical interventions or gene therapies in order to ameliorate phenotypic manifestations, and expect this will bolster our understanding of the biological systems and mechanisms underlying these and other disorders. Furthermore, the knowledge we gain from sequencing will enhance genetic counseling, carrier screening, and prenatal testing, and perhaps even lead to phenotypic expansion for known disorders. Our efforts in Mendelian genetics have been tremendously exciting and informative; we look forward to the prospects the future will hold.

Bibliography

1000 Genomes Project Consortium, G. A., Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., … McVean, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65. https://doi.org/10.1038/nature11632

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., … Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Agassandian, M., & Mallampalli, R. K. (2013). Surfactant phospholipid metabolism. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1831*(3), 612–625. https://doi.org/10.1016/j.bbalip.2012.09.010

Aitchison, A. J., Arsenault, D. J., & Ridgway, N. D. (2015). Nuclear-localized CTP:phosphocholine cytidylyltransferase regulates phosphatidylcholine synthesis required for lipid droplet biogenesis. *Molecular Biology of the Cell*, *26*(16), 2927–2938. https://doi.org/10.1091/mbc.E15-03-0159

Al-Jassar, C., Knowles, T., Jeeves, M., Kami, K., Behr, E., Bikker, H., … Chidgey, M. (2011). The nonlinear structure of the desmoplakin plakin domain and the effects of cardiomyopathy-linked mutations. *Journal of Molecular Biology*, *411*(5), 1049–61. https://doi.org/10.1016/j.jmb.2011.06.047

Angelis, A., Tordrup, D., & Kanavos, P. (2015). Socio-economic burden of rare diseases: A systematic review of cost of illness evidence. *Health Policy*, *119*(7), 964–979. https://doi.org/10.1016/J.HEALTHPOL.2014.12.016

Astejada, M. N., Goto, K., Nagano, A., Ura, S., Noguchi, S., Nonaka, I., … Hayashi, Y. K. (2007). Emerinopathy and laminopathy clinical, pathological and molecular features of muscular dystrophy with nuclear envelopathy in Japan. *Acta Myologica : Myopathies and Cardiomyopathies : Official Journal of the Mediterranean Society of Myology*, *26*(3), 159–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18646565

Barber, T., Esteban-Pretel, G., Marín, M., & Timoneda, J. (2014). Vitamin A Deficiency and Alterations in the Extracellular Matrix. *Nutrients*, *6*(11), 4984–5017. https://doi.org/10.3390/nu6114984

Barbour, S. E., Kapur, A., & Deal, C. L. (1999). Regulation of phosphatidylcholine homeostasis by calcium-independent phospholipase A2. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1439*(1), 77–88. https://doi.org/10.1016/S1388-1981(99)00078-5

Bean, L. J. H., Tinker, S. W., da Silva, C., & Hegde, M. R. (2013). Free the data: one laboratory's approach to knowledge-based genomic variant classification and preparation for EMR integration of genomic data. *Human Mutation*, *34*(9), 1183–8. https://doi.org/10.1002/humu.22364

Bell, C. J., Dinwiddie, D. L., Miller, N. A., Hateley, S. L., Ganusova, E. E., Mudge, J., … Kingsmore, S. F. (2011). Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing. *Science Translational Medicine*, *3*(65), 65ra4-65ra4. https://doi.org/10.1126/scitranslmed.3001756

Braverman, N., Chen, L., Lin, P., Obie, C., Steel, G., Douglas, P., … Valle, D. (2002).

Mutation analysis of PEX7 in 60 probands with rhizomelic chondrodysplasia punctata and functional correlations of genotype with phenotype. *Human Mutation*, *20*(4), 284–97. https://doi.org/10.1002/humu.10124

Bridges, J. P., Ikegami, M., Brilli, L. L., Chen, X., Mason, R. J., & Shannon, J. M. (2010). LPCAT1 regulates surfactant phospholipid synthesis and is required for transitioning to air breathing in mice. *The Journal of Clinical Investigation*, *120*(5), 1736–48. https://doi.org/10.1172/JCI38061

Carmignac, V., Salih, M. A. M., Quijano-Roy, S., Marchand, S., Al Rayess, M. M., Mukhtar, M. M., … Ferreiro, A. (2007). C-terminal titin deletions cause a novel early-onset myopathy with fatal cardiomyopathy. *Annals of Neurology*, *61*(4), 340–51. https://doi.org/10.1002/ana.21089

Carter, J. M., Demizieux, L., Campenot, R. B., Vance, D. E., & Vance, J. E. (2008). Phosphatidylcholine biosynthesis via CTP:phosphocholine cytidylyltransferase 2 facilitates neurite outgrowth and branching. *The Journal of Biological Chemistry*, *283*(1), 202–12. https://doi.org/10.1074/jbc.M706531200

Cavallone, L., Arcand, S. L., Maugard, C. M., Nolet, S., Gaboury, L. A., Mes-Masson, A.-M., … Tonin, P. N. (2010). Comprehensive BRCA1 and BRCA2 mutation analyses and review of French Canadian families with at least three cases of breast cancer. *Familial Cancer*, *9*(4), 507–17. https://doi.org/10.1007/s10689-010-9372-3

Chong, J. X., Buckingham, K. J., Jhangiani, S. N., Boehm, C., Sobreira, N., Smith, J. D., … Bamshad, M. J. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *The American Journal of Human*

Genetics, *97*(2), 199–215. https://doi.org/10.1016/J.AJHG.2015.06.009

Cornell, R. B., & Ridgway, N. D. (2015). CTP:phosphocholine cytidylyltransferase: Function, regulation, and structure of an amphitropic enzyme required for membrane biogenesis. *Progress in Lipid Research*, *59*, 147–171. https://doi.org/10.1016/j.plipres.2015.07.001

Cornell, R., & Vance, D. E. (1987). Translocation of CTP:phosphocholine cytidylyltransferase from cytosol to membranes in HeLa cells: stimulation by fatty acid, fatty alcohol, mono- and diacylglycerol. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, *919*(1), 26–36. https://doi.org/10.1016/0005-2760(87)90214-1

Cowan, J., Li, D., Gonzalez-Quintana, J., Morales, A., & Hershberger, R. E. (2010). Morphological Analysis of 13 LMNA Variants Identified in a Cohort of 324 Unrelated Patients With Idiopathic or Familial Dilated Cardiomyopathy. *Circulation: Cardiovascular Genetics*, *3*(1), 6–14. https://doi.org/10.1161/CIRCGENETICS.109.905422

Cremers, F. P., van de Pol, D. J., van Driel, M., den Hollander, A. I., van Haren, F. J., Knoers, N. V, … Hoyng, C. B. (1998). Autosomal recessive retinitis pigmentosa and cone-rod dystrophy caused by splice site mutations in the Stargardt's disease gene ABCR. *Human Molecular Genetics*, *7*(3), 355–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9466990

Cui, Z., Houweling, M., Chen, M. H., Record, M., Chap, H., Vance, D. E., & Tercé, F. (1996). A genetic defect in phosphatidylcholine biosynthesis triggers apoptosis in

Chinese hamster ovary cells. *The Journal of Biological Chemistry*, *271*(25), 14668–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8663247

Davydov, E. V, Goode, D. L., Sirota, M., Cooper, G. M., Sidow, A., & Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Computational Biology*, *6*(12), e1001025. https://doi.org/10.1371/journal.pcbi.1001025

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J. R., Hartl, C., … Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. https://doi.org/10.1038/ng.806

Ding, Z., Taneva, S. G., Huang, H. K. H., Campbell, S. A., Semenec, L., Chen, N., & Cornell, R. B. (2012). A 22-mer Segment in the Structurally Pliable Regulatory Domain of Metazoan CTP: Phosphocholine Cytidylyltransferase Facilitates Both Silencing and Activating Functions. *Journal of Biological Chemistry*, *287*(46), 38980–38991. https://doi.org/10.1074/jbc.M112.402081

Dorschner, M. O., Amendola, L. M., Turner, E. H., Robertson, P. D., Shirts, B. H., Gallego, C. J., … Jarvik, G. P. (2013). Actionable, pathogenic incidental findings in 1,000 participants' exomes. *American Journal of Human Genetics*, *93*(4), 631–40. https://doi.org/10.1016/j.ajhg.2013.08.006

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., … Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Edwards, S. M., Kote-Jarai, Z., Hamoudi, R., & Eeles, R. A. (2001). An improved high throughput heteroduplex mutation detection system for screening BRCA2 mutations-fluorescent mutation detection (F-MD). *Human Mutation*, *17*(3), 220–32. https://doi.org/10.1002/humu.7

Ehara, S., Kim, O. H., Maisawa, S., Takasago, Y., & Nishimura, G. (1997). Axial spondylometaphyseal dysplasia. *European Journal of Pediatrics*, *156*(8), 627–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9266195

Ehlermann, P., Weichenhan, D., Zehelein, J., Steen, H., Pribe, R., Zeller, R., … Katus, H. A. (2008). Adverse events in families with hypertrophic or dilated cardiomyopathy and mutations in the MYBPC3gene. *BMC Medical Genetics*, *9*(1), 95. https://doi.org/10.1186/1471-2350-9-95

Fagone, P., & Jackowski, S. (2013). Phosphatidylcholine and the CDP–choline cycle. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1831*(3), 523–532. https://doi.org/10.1016/j.bbalip.2012.09.009

Findlay, G. M., Boyle, E. A., Hause, R. J., Klein, J. C., & Shendure, J. (2014). Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*, *513*(7516), 120–123. https://doi.org/10.1038/nature13695

Friedman, J. S., Chang, B., Krauth, D. S., Lopez, I., Waseem, N. H., Hurd, R. E., … Swaroop, A. (2010). Loss of lysophosphatidylcholine acyltransferase 1 leads to photoreceptor degeneration in rd11 mice. *Proceedings of the National Academy of Sciences*, *107*(35), 15523–15528. https://doi.org/10.1073/pnas.1002897107

Garcia Segarra, N., Mittaz, L., Campos-Xavier, A. B., Bartels, C. F., Tuysuz, B., Alanay,

Y., … Bonafé, L. (2012). The diagnostic challenge of progressive pseudorheumatoid dysplasia (PPRD): a review of clinical features, radiographic features, and WISP3 mutations in 63 affected individuals. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *160C*(3), 217–29. https://doi.org/10.1002/ajmg.c.31333

Gehrig, K., Cornell, R. B., & Ridgway, N. D. (2008). Expansion of the Nucleoplasmic Reticulum Requires the Coordinated Activity of Lamins and CTP:Phosphocholine Cytidylyltransferase. *Molecular Biology of the Cell*, *19*(1), 237–247. https://doi.org/10.1091/mbc.E07-02-0179

Gerull, B., Gramlich, M., Atherton, J., McNabb, M., Trombitás, K., Sasse-Klaassen, S., … Thierfelder, L. (2002). Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nature Genetics*, *30*(2), 201–4. https://doi.org/10.1038/ng815

Glunde, K., Bhujwalla, Z. M., & Ronen, S. M. (2011). Choline metabolism in malignant transformation. *Nature Reviews Cancer*, *11*(12), 835–48. https://doi.org/10.1038/nrc3162

Goldstein, J. L., & Brown, M. S. (2015). A century of cholesterol and coronaries: from plaques to genes to statins. *Cell*, *161*(1), 161–72. https://doi.org/10.1016/j.cell.2015.01.036

Goss, V., Hunt, A. N., & Postle, A. D. (2013). Regulation of lung surfactant phospholipid synthesis and metabolism. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1831*(2), 448–458.

https://doi.org/10.1016/j.bbalip.2012.11.009

Green, R. C., Berg, J. S., Grody, W. W., Kalia, S. S., Korf, B. R., Martin, C. L., … American College of Medical Genetics and Genomics. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genetics in Medicine*, *15*(7), 565–574. https://doi.org/10.1038/gim.2013.73

Guo, Y., Walther, T. C., Rao, M., Stuurman, N., Goshima, G., Terayama, K., … Farese, R. V. (2008). Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. *Nature*, *453*(7195), 657–661. https://doi.org/10.1038/nature06928

Gussow, A. B., Petrovski, S., Wang, Q., Allen, A. S., & Goldstein, D. B. (2016). The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. *Genome Biology*, *17*(1), 9. https://doi.org/10.1186/s13059-016-0869-4

Hackman, P., Vihola, A., Haravuori, H., Marchand, S., Sarparanta, J., De Seze, J., … Udd, B. (2002). Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. *American Journal of Human Genetics*, *71*(3), 492–500. https://doi.org/10.1086/342380

Hamel, C. P. (2007). Cone rod dystrophies. *Orphanet Journal of Rare Diseases*, *2*(1), 7. https://doi.org/10.1186/1750-1172-2-7

Hamosh, A., Sobreira, N., Hoover-Fong, J., Sutton, V. R., Boehm, C., Schiettecatte, F., & Valle, D. (2013). PhenoDB: a new web-based tool for the collection, storage, and

analysis of phenotypic features. *Human Mutation*, *34*(4), 566–71. https://doi.org/10.1002/humu.22283

Herman, D. S., Lam, L., Taylor, M. R. G., Wang, L., Teekakirikul, P., Christodoulou, D., … Seidman, C. E. (2012). Truncations of Titin Causing Dilated Cardiomyopathy. *New England Journal of Medicine*, *366*(7), 619–628. https://doi.org/10.1056/NEJMoa1110186

Herman, G. E., Kelley, R. I., Pureza, V., Smith, D., Kopacz, K., Pitt, J., … Metzenberg, A. B. (n.d.). Characterization of mutations in 22 females with X-linked dominant chondrodysplasia punctata (Happle syndrome). *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *4*(6), 434–8. https://doi.org/10.109700125817-200211000-00006

Hoover-Fong, J., Sobreira, N., Jurgens, J., Modaff, P., Blout, C., Moser, A., … Pauli, R. M. (2014). Mutations in PCYT1A, Encoding a Key Regulator of Phosphatidylcholine Metabolism, Cause Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy. *The American Journal of Human Genetics*, *94*(1), 105–112. https://doi.org/10.1016/j.ajhg.2013.11.018

Hörl, G., Wagner, A., Cole, L. K., Malli, R., Reicher, H., Kotzbeck, P., … Steyrer, E. (2011). Sequential synthesis and methylation of phosphatidylethanolamine promote lipid droplet biosynthesis and stability in tissue culture and in vivo. *The Journal of Biological Chemistry*, *286*(19), 17338–50. https://doi.org/10.1074/jbc.M111.234534

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., … Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature*

Biotechnology, *31*(9), 827–32. https://doi.org/10.1038/nbt.2647

Hubbard, T. J. P., Aken, B. L., Ayling, S., Ballester, B., Beal, K., Bragin, E., … Flicek, P. (2009). Ensembl 2009. *Nucleic Acids Research*, *37*(Database), D690–D697. https://doi.org/10.1093/nar/gkn828

Hurvitz, J. R., Suwairi, W. M., Van Hul, W., El-Shanti, H., Superti-Furga, A., Roudier, J., … Warman, M. L. (1999). Mutations in the CCN gene family member WISP3 cause progressive pseudorheumatoid dysplasia. *Nature Genetics*, *23*(1), 94–98. https://doi.org/10.1038/12699

Isidor, B., Le Merrer, M., Ramos, E., Baron, S., & David, A. (2009). Cone-rod dystrophy, growth hormone deficiency and spondyloepiphyseal dysplasia: report of a new case without nephronophtisis. *American Journal of Medical Genetics. Part A*, *149A*(4), 788–92. https://doi.org/10.1002/ajmg.a.32343

Itoh-Satoh, M., Hayashi, T., Nishi, H., Koga, Y., Arimura, T., Koyanagi, T., … Kimura, A. (2002). Titin Mutations as the Molecular Basis for Dilated Cardiomyopathy. *Biochemical and Biophysical Research Communications*, *291*(2), 385–393. https://doi.org/10.1006/bbrc.2002.6448

Jääskeläinen, P., Kuusisto, J., Miettinen, R., Kärkkäinen, P., Kärkkäinen, S., Heikkinen, S., … Laakso, M. (2002). Mutations in the cardiac myosin-binding protein C gene are the predominant cause of familial hypertrophic cardiomyopathy in eastern Finland. *Journal of Molecular Medicine (Berlin, Germany)*, *80*(7), 412–22. https://doi.org/10.1007/s00109-002-0323-9

Jackowski, S., Rehg, J. E., Zhang, Y.-M., Wang, J., Miller, K., Jackson, P., & Karim, M.

A. (2004). Disruption of CCT 2 Expression Leads to Gonadal Dysfunction. *Molecular and Cellular Biology*, *24*(11), 4720–4733. https://doi.org/10.1128/MCB.24.11.4720-4733.2004

Jacobs, R. L., Devlin, C., Tabas, I., & Vance, D. E. (2004). Targeted deletion of hepatic CTP:phosphocholine cytidylyltransferase alpha in mice decreases plasma high density and very low density lipoproteins. *The Journal of Biological Chemistry*, *279*(45), 47402–10. https://doi.org/10.1074/jbc.M404027200

Jarvik, G. P., Amendola, L. M., Berg, J. S., Brothers, K., Clayton, E. W., Chung, W., … Sun, K. (2014). Return of genomic results to research participants: the floor, the ceiling, and the choices in between. *American Journal of Human Genetics*, *94*(6), 818–26. https://doi.org/10.1016/j.ajhg.2014.04.009

Johnston, J. J., Rubinstein, W. S., Facio, F. M., Ng, D., Singh, L. N., Teer, J. K., … Biesecker, L. G. (2012). Secondary variants in individuals undergoing exome sequencing: screening of 572 individuals identifies high-penetrance mutations in cancer-susceptibility genes. *American Journal of Human Genetics*, *91*(1), 97–108. https://doi.org/10.1016/j.ajhg.2012.05.021

Judkins, T., Hendrickson, B. C., Deffenbaugh, A. M., Eliason, K., Leclair, B., Norton, M. J., … Scholl, T. (2005). Application of Embryonic Lethal or Other Obvious Phenotypes to Characterize the Clinical Significance of Genetic Variants Found in *Trans* with Known Deleterious Mutations. *Cancer Research*, *65*(21), 10096–10103. https://doi.org/10.1158/0008-5472.CAN-05-1241

Karim, M., Jackson, P., & Jackowski, S. (2003). Gene structure, expression and

identification of a new CTP:phosphocholine cytidylyltransferase beta isoform. *Biochimica et Biophysica Acta*, *1633*(1), 1–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12842190

Keenan, T. W., & Morré, D. J. (1970). Phospholipid class and fatty acid composition of golgi apparatus isolated from rat liver and comparison with other cell fractions. *Biochemistry*, *9*(1), 19–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4312390

KENNEDY, E. P., & WEISS, S. B. (1956). The function of cytidine coenzymes in the biosynthesis of phospholipides. *The Journal of Biological Chemistry*, *222*(1), 193–214. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13366993

Kent, C. (2005). Regulatory enzymes of phosphatidylcholine biosynthesis: a personal perspective. *Biochimica et Biophysica Acta*, *1733*(1), 53–66. https://doi.org/10.1016/j.bbalip.2004.12.008

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315. https://doi.org/10.1038/ng.2892

Kitoh, H., Kaneko, H., Kondo, M., Yamamoto, T., Ishiguro, N., & Nishimura, G. (2011). Spondylometaphyseal dysplasia with cone-rod dystrophy. *American Journal of Medical Genetics Part A*, *155*(4), 845–849. https://doi.org/10.1002/ajmg.a.33898

Kitos, T. E., Drobnies, A., Ng, M. N. P., Wen, Y., & Cornell, R. B. (2006). Contribution of lipid mediators to the regulation of phosphatidylcholine synthesis by angiotensin. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*,

1761(2), 261–271. https://doi.org/10.1016/j.bbalip.2006.02.015

Krahmer, N., Guo, Y., Wilfling, F., Hilger, M., Lingrell, S., Heger, K., … Walther, T. C. (2011). Phosphatidylcholine Synthesis for Lipid Droplet Expansion Is Mediated by Localized Activation of CTP:Phosphocholine Cytidylyltransferase. *Cell Metabolism*, *14*(4), 504–515. https://doi.org/10.1016/j.cmet.2011.07.013

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(8), 1073–1081. https://doi.org/10.1038/nprot.2009.86

Lagace, T. A., & Ridgway, N. D. (2005). The Rate-limiting Enzyme in Phosphatidylcholine Synthesis Regulates Proliferation of the Nucleoplasmic Reticulum. *Molecular Biology of the Cell*, *16*(3), 1120–1130. https://doi.org/10.1091/mbc.E04-10-0874

Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott, D. R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research*, *42*(Database issue), D980-5. https://doi.org/10.1093/nar/gkt1113

Lands, W. E. M. (n.d.). Metabolism of Glycerolipides: A comparison of lecithin and triglyceride synthesis. Retrieved from: http://www.jbc.org.ezp.welch.jhmi.edu/content/231/2/883.full.pdf

Lange, S., Xiang, F., Yakovenko, A., Vihola, A., Hackman, P., Rostkova, E., … Gautel, M. (2005). The Kinase Domain of Titin Controls Muscle Gene Expression and Protein Turnover. *Science*, *308*(5728), 1599–1603.

https://doi.org/10.1126/science.1110463

Lanske, B., Karaplis, A. C., Lee, K., Luz, A., Vortkamp, A., Pirro, A., … Kronenberg, H. M. (1996). PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth. *Science (New York, N.Y.)*, *273*(5275), 663–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8662561

LaVail, M. M. (1976). Rod outer segment disk shedding in rat retina: relationship to cyclic lighting. *Science (New York, N.Y.)*, *194*(4269), 1071–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/982063

Lawrence, L., Sincan, M., Markello, T., Adams, D. R., Gill, F., Godfrey, R., … Boerkoel, C. F. (2014). The implications of familial incidental findings from exome sequencing: the NIH Undiagnosed Diseases Program experience. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *16*(10), 741–50. https://doi.org/10.1038/gim.2014.29

Lee, J., Johnson, J., Ding, Z., Paetzel, M., & Cornell, R. B. (2009). Crystal Structure of a Mammalian CTP: Phosphocholine Cytidylyltransferase Catalytic Domain Reveals Novel Active Site Residues within a Highly Conserved Nucleotidyltransferase Fold. *Journal of Biological Chemistry*, *284*(48), 33535–33548. https://doi.org/10.1074/jbc.M109.053363

Lee, M., Roos, P., Sharma, N., Atalar, M., Evans, T. A., Pellicore, M. J., … Cutting, G. R. (2017). Systematic Computational Identification of Variants That Activate Exonic and Intronic Cryptic Splice Sites. *American Journal of Human Genetics*, *100*(5), 751–765. https://doi.org/10.1016/j.ajhg.2017.04.001

Lee, S., Abecasis, G. R., Boehnke, M., & Lin, X. (2014). Rare-variant association analysis: study designs and statistical tests. *American Journal of Human Genetics*, *95*(1), 5–23. https://doi.org/10.1016/j.ajhg.2014.06.009

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., … Consortium, E. A. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291. https://doi.org/10.1038/nature19057

Lelieveld, S. H., Wiel, L., Venselaar, H., Pfundt, R., Vriend, G., Veltman, J. A., … Gilissen, C. (2017). Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *The American Journal of Human Genetics*, *101*(3), 478–484. https://doi.org/10.1016/j.ajhg.2017.08.004

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Z., Agellon, L. B., Allen, T. M., Umeda, M., Jewell, L., Mason, A., & Vance, D. E. (2006). The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. *Cell Metabolism*, *3*(5), 321–331. https://doi.org/10.1016/J.CMET.2006.03.007

Lloyd, K. C. K. (2011). A knockout mouse resource for the biomedical research

community. *Annals of the New York Academy of Sciences*, *1245*, 24–6. https://doi.org/10.1111/j.1749-6632.2011.06311.x

Lopes, L. R., Zekavati, A., Syrris, P., Hubank, M., Giambartolomei, C., Dalageorgou, C., … Elliott, P. M. (2013). Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. *Journal of Medical Genetics*, *50*(4), 228–39. https://doi.org/10.1136/jmedgenet-2012-101270

Lu, C., Jain, S. U., Hoelper, D., Bechet, D., Molden, R. C., Ran, L., … Lewis, P. W. (2016). Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science (New York, N.Y.)*, *352*(6287), 844–9. https://doi.org/10.1126/science.aac7272

Malone, K. E., Daling, J. R., Neal, C., Suter, N. M., O'Brien, C., Cushing-Haugen, K., … Ostrander, E. A. (2000). Frequency of BRCA1/BRCA2 mutations in a population-based sample of young breast carcinoma cases. *Cancer*, *88*(6), 1393–402. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10717622

Marik, I., Marikova, O., Zemkova, D., Kuklik, M., & Kozlowski, K. (2004). Dominantly inherited progressive pseudorheumatoid dysplasia with hypoplastic toes. *Skeletal Radiology*, *33*(3), 157–64. https://doi.org/10.1007/s00256-003-0708-z

Matsushita, M., Kitoh, H., Ohkawara, B., Mishima, K., Kaneko, H., Ito, M., … Ohno, K. (2013). Meclozine Facilitates Proliferation and Differentiation of Chondrocytes by Attenuating Abnormally Activated FGFR3 Signaling in Achondroplasia. *PLoS ONE*, *8*(12), e81569. https://doi.org/10.1371/journal.pone.0081569

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., …

DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–303. https://doi.org/10.1101/gr.107524.110

Mitsuhashi, S., Ohkuma, A., Talim, B., Karahashi, M., Koumura, T., Aoyama, C., … Nishino, I. (2011). A congenital muscular dystrophy with mitochondrial structural abnormalities caused by defective de novo phosphatidylcholine biosynthesis. *American Journal of Human Genetics*, *88*(6), 845–851. https://doi.org/10.1016/j.ajhg.2011.05.010

Moolman-Smook, J. C., De Lange, W. J., Bruwer, E. C., Brink, P. A., & Corfield, V. A. (1999). The origins of hypertrophic cardiomyopathy-causing mutations in two South African subpopulations: a unique profile of both independent and founder events. *American Journal of Human Genetics*, *65*(5), 1308–20. https://doi.org/10.1086/302623

Morak, M., Massdorf, T., Sykora, H., Kerscher, M., & Holinski-Feder, E. (2011). First evidence for digenic inheritance in hereditary colorectal cancer by mutations in the base excision repair genes. *European Journal of Cancer (Oxford, England : 1990)*, *47*(7), 1046–55. https://doi.org/10.1016/j.ejca.2010.11.016

Moyer, T. C., & Holland, A. J. (2015). Generation of a conditional analog-sensitive kinase in human cells using CRISPR/Cas9-mediated genome engineering. In *Methods in cell biology* (Vol. 129, pp. 19–36). https://doi.org/10.1016/bs.mcb.2015.03.017

Ng, M. N. P., Kitos, T. E., & Cornell, R. B. (2004). Contribution of lipid second

messengers to the regulation of phosphatidylcholine synthesis during cell cycle re-entry. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1686*(1–2), 85–99. https://doi.org/10.1016/j.bbalip.2004.09.001

Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M., … Bamshad, M. J. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nature Genetics*, *42*(1), 30–35. https://doi.org/10.1038/ng.499

Nishimura, G., Saitoh, Y., Okuzumi, S., Imaizumi, K., Hayasaka, K., & Hashimoto, M. (1998). Spondyloepiphyseal dysplasia with accumulation of glycoprotein in the chondrocytes: spondyloepiphyseal dysplasia, Stanescu type. *Skeletal Radiology*, *27*(4), 188–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9592900

Olivotto, I., Girolami, F., Ackerman, M. J., Nistri, S., Bos, J. M., Zachara, E., … Cecchi, F. (2008). Myofilament Protein Gene Mutation Screening and Outcome of Patients With Hypertrophic Cardiomyopathy. *Mayo Clinic Proceedings*, *83*(6), 630–638. https://doi.org/10.4065/83.6.630

Parks, S. B., Kushner, J. D., Nauman, D., Burgess, D., Ludwigsen, S., Peterson, A., … Hershberger, R. E. (2008). Lamin A/C mutation analysis in a cohort of 324 unrelated patients with idiopathic or familial dilated cardiomyopathy. *American Heart Journal*, *156*(1), 161–169. https://doi.org/10.1016/j.ahj.2008.01.026

Payne, F., Lim, K., Girousse, A., Brown, R. J., Kory, N., Robbins, A., … Savage, D. B. (2014). Mutations disrupting the Kennedy phosphatidylcholine pathway in humans with congenital lipodystrophy and fatty liver disease. *Proceedings of the National Academy of Sciences*, *111*(24), 8901–8906.

https://doi.org/10.1073/pnas.1408523111

Payne, S. R., & Kemp, C. J. (2005). Tumor suppressor genetics. *Carcinogenesis*, *26*(12), 2031–2045. https://doi.org/10.1093/carcin/bgi223

Pelechs, S. L., Pritchard, P. H., Brindleyg, D. N., & Vance, D. E. (1983). Fatty Acids Promote Translocation of CTP:Phosphocholine Cytidylyltransferase to the Endoplasmic Reticulkm and Stimulate Rat Hepatic Phosphatidylcholine Synthesis. *The Journal of Biological Chemistry*, *258*(11), 6782–6788. Retrieved from http://www.jbc.org.ezp.welch.jhmi.edu/content/258/11/6782.full.pdf

Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S., & Goldstein, D. B. (2013). Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes. *PLoS Genetics*, *9*(8), e1003709. https://doi.org/10.1371/journal.pgen.1003709

Phelan, C. M., Dapic, V., Tice, B., Favis, R., Kwan, E., Barany, F., … Monteiro, A. N. A. (2005). Classification of BRCA1 missense variants of unknown clinical significance. *Journal of Medical Genetics*, *42*(2), 138–146. https://doi.org/10.1136/jmg.2004.024711

Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2004). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, *33*(Database issue), D501–D504. https://doi.org/10.1093/nar/gki025

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., … EUROGENE Heart Failure Project. (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation*, *107*(17), 2227–32. https://doi.org/10.1161/01.CIR.0000066323.15244.54

Ridsdale, R., Tseu, I., Wang, J., & Post, M. (2001). CTP:phosphocholine cytidylyltransferase alpha is a cytosolic protein in pulmonary epithelial cells and tissues. *The Journal of Biological Chemistry*, *276*(52), 49148–55. https://doi.org/10.1074/jbc.M103566200

Rodríguez-García, M. I., Monserrat, L., Ortiz, M., Fernández, X., Cazón, L., Núñez, L., … Hermida-Prieto, M. (2010). Screening mutations in myosin binding protein C3 gene in a cohort of patients with Hypertrophic Cardiomyopathy. *BMC Medical Genetics*, *11*(1), 67. https://doi.org/10.1186/1471-2350-11-67

Rukavina, I., Mortier, G., Van Laer, L., Frković, M., Đapić, T., & Jelušić, M. (2014). Mutation in the type II collagen gene (COL2AI) as a cause of primary osteoarthritis associated with mild spondyloepiphyseal involvement. *Seminars in Arthritis and Rheumatism*, *44*(1), 101–4. https://doi.org/10.1016/j.semarthrit.2014.03.003

Saleheen, D., Natarajan, P., Armean, I. M., Zhao, W., Rasheed, A., Khetarpal, S. A., … Kathiresan, S. (2017). Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature*, *544*(7649), 235–239. https://doi.org/10.1038/nature22034

Samowitz, W. S., Curtin, K., Lin, H. H., Robertson, M. A., Schaffer, D., Nichols, M., …

Slattery, M. L. (2001). The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology*, *121*(4), 830–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11606497

Satoh, M., Takahashi, M., Sakamoto, T., Hiroe, M., Marumo, F., & Kimura, A. (1999). Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochemical and Biophysical Research Communications*, *262*(2), 411–7. https://doi.org/10.1006/bbrc.1999.1221

Sen, M., Cheng, Y.-H., Goldring, M. B., Lotz, M. K., & Carson, D. A. (2004). WISP3-dependent regulation of type II collagen and aggrecan production in chondrocytes. *Arthritis and Rheumatism*, *50*(2), 488–97. https://doi.org/10.1002/art.20005

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, *29*(1), 308–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11125122

Sieber, O. M., Lipton, L., Crabtree, M., Heinimann, K., Fidalgo, P., Phillips, R. K. S., … Tomlinson, I. P. M. (2003). Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations in *MYH*. *New England Journal of Medicine*, *348*(9), 791–799. https://doi.org/10.1056/NEJMoa025283

Smedley, D., Schubach, M., Jacobsen, J. O. B., Köhler, S., Zemojtel, T., Spielmann, M., … Robinson, P. N. (2016). A Whole-Genome Analysis Framework for Effective Identification of Pathogenic Regulatory Variants in Mendelian Disease. *The American Journal of Human Genetics*, *99*(3), 595–606.

https://doi.org/10.1016/j.ajhg.2016.07.005

Sobreira, N., Schiettecatte, F., Boehm, C., Valle, D., & Hamosh, A. (2015). New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. *Human Mutation*, *36*(4), 425–31. https://doi.org/10.1002/humu.22769

Sobreira, N., Schiettecatte, F., Valle, D., & Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Human Mutation*, *36*(10), 928–30. https://doi.org/10.1002/humu.22844

Sousa, S. B., Russell-Eggitt, I., Hall, C., Hall, B. D., & Hennekam, R. C. M. (2008). Further delineation of spondylometaphyseal dysplasia with cone-rod dystrophy. *American Journal of Medical Genetics. Part A*, *146A*(24), 3186–94. https://doi.org/10.1002/ajmg.a.32576

Stenson, P. D., Ball, E. V, Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S. T., … Cooper, D. N. (2003). Human Gene Mutation Database (HGMD): 2003 update. *Human Mutation*, *21*(6), 577–81. https://doi.org/10.1002/humu.10212

Sweitzer, T. D., & Kent, C. (1994). Expression of Wild-Type and Mutant Rat Liver CTP: Phosphocholine Cytidylyltransferase in a Cytidylyltransferase-Deficient Chinese Hamster Ovary Cell Line. *Archives of Biochemistry and Biophysics*, *311*(1), 107–116. https://doi.org/10.1006/ABBI.1994.1215

Tabe, S., Hikiji, H., Ariyoshi, W., Hashidate-Yoshida, T., Shindou, H., Shimizu, T., … Nishihara, T. (2017). Lysophosphatidylcholine acyltransferase 4 is involved in chondrogenic differentiation of ATDC5 cells. *Scientific Reports*, *7*(1), 16701.

https://doi.org/10.1038/s41598-017-16902-4

Tavtigian, S. V, Deffenbaugh, A. M., Yin, L., Judkins, T., Scholl, T., Samollow, P. B., … Thomas, A. (2006). Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *Journal of Medical Genetics*, *43*(4), 295–305. https://doi.org/10.1136/jmg.2005.033878

Terry, S. F. (2015). Obama's Precision Medicine Initiative. *Genetic Testing and Molecular Biomarkers*, *19*(3), 113–4. https://doi.org/10.1089/gtmb.2015.1563

Testa, F., Filippelli, M., Brunetti-Pierri, R., Di Fruscio, G., Di Iorio, V., Pizzo, M., … Banfi, S. (2017). Mutations in the PCYT1A gene are responsible for isolated forms of retinal dystrophy. *European Journal of Human Genetics*, *25*(5), 651–655. https://doi.org/10.1038/ejhg.2017.23

Tian, Y., Zhou, R., Rehg, J. E., & Jackowski, S. (2007). Role of Phosphocholine Cytidylyltransferase in Lung Development. *Molecular and Cellular Biology*, *27*(3), 975–982. https://doi.org/10.1128/MCB.01512-06

Turell, M., Morrison, S., & Traboulsi, E. I. (2010). Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy. *Ophthalmic Genetics*, *31*(1), 12–17. https://doi.org/10.3109/13816810903397812

Vance, J. E., & Vance, D. E. (2004). Phospholipid biosynthesis in mammalian cells. *Biochemistry and Cell Biology*, *82*(1), 113–128. https://doi.org/10.1139/o03-073

Walkey$, C. J., Kalmar, G. B., & Cornel15, R. B. (1994). Overexpression of Rat Liver CTP:Phosphocholine Cytidylyltransferase Accelerates Phosphatidylcholine

Synthesis and Degradation. *The Journal of Biological Chemistry*, *269*(8), 5742–5749. Retrieved from http://www.jbc.org.ezp.welch.jhmi.edu/content/269/8/5742.full.pdf?sid=acb5babb-0277-4f08-a1df-481f8c6b635e

Walters, B. A., Raff, M. L., Hoeve, J. Ver, Tesser, R., Langer, L. O., France, T. D., … Pauli, R. M. (2004). Spondylometaphyseal dysplasia with cone-rod dystrophy. *American Journal of Medical Genetics. Part A*, *129A*(3), 265–76. https://doi.org/10.1002/ajmg.a.30145

Walther, T. C., Chung, J., & Farese, R. V. (2017). Lipid Droplet Biogenesis. *Annual Review of Cell and Developmental Biology*, *33*(1), 491–510. https://doi.org/10.1146/annurev-cellbio-100616-060608

Wang, L., Guo, D., Cao, J., Gong, L., Kamm, K. E., Regalado, E., … Milewicz, D. M. (2010). Mutations in Myosin Light Chain Kinase Cause Familial Aortic Dissections. *The American Journal of Human Genetics*, *87*(5), 701–707. https://doi.org/10.1016/j.ajhg.2010.10.006

Wang, L., Magdaleno, S., Tabas, I., & Jackowski, S. (2005). Early Embryonic Lethality in Mice with Targeted Deletion of the CTP:Phosphocholine Cytidylyltransferase Gene (Pcyt1a). *Molecular and Cellular Biology*, *25*(8), 3357–3363. https://doi.org/10.1128/MCB.25.8.3357-3363.2005

Wang, M., Man, X.-F., Liu, Y.-Q., Liao, E.-Y., Shen, Z.-F., Luo, X.-H., … Zhou, H.-D. (2013). Dysfunction of collagen synthesis and secretion in chondrocytes induced by wisp3 mutation. *International Journal of Endocrinology*, *2013*, 679763. https://doi.org/10.1155/2013/679763

Wang, Y., Macdonald, J. I. S., & Kent4, C. (1993). Regulation of CTP:Phosphocholine Cytidylyltransferase in HeLa Cells: Effect of oleate on phosphorylation and intracellular localization. *The Journal of Biological Chemistry, 268*(8), 5512–5518. Retrieved from http://www.jbc.org.ezp.welch.jhmi.edu/content/268/8/5512.full.pdf

Warman, M. L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., LeMerrer, M., … Superti-Furga, A. (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. *American Journal of Medical Genetics Part A*, *155*(5), 943–968. https://doi.org/10.1002/ajmg.a.33909

Wiel, L., Venselaar, H., Veltman, J. A., Vriend, G., & Gilissen, C. (2017). Aggregation of population-based genetic variation over protein domain homologues and its potential use in genetic diagnostics. *Human Mutation*, *38*(11), 1454–1463. https://doi.org/10.1002/humu.23313

Yamamoto, G. L., Baratela, W. A. R., Almeida, T. F., Lazar, M., Afonso, C. L., Oyamada, M. K., … Bertola, D. R. (2014). Mutations in PCYT1A Cause Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy. *The American Journal of Human Genetics*, *94*(1), 113–119. https://doi.org/10.1016/j.ajhg.2013.11.022

Yang, Z., Chen, Y., Lillo, C., Chien, J., Yu, Z., Michaelides, M., … Zhang, K. (2008). Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice. *The Journal of Clinical Investigation*, *118*(8), 2908–16. https://doi.org/10.1172/JCI35891

Zhang, D., Tang, W., Yao, P. M., Yang, C., Xie, B., Jackowski, S., & Tabas, I. (2000). Macrophages deficient in CTP:Phosphocholine cytidylyltransferase-alpha are viable

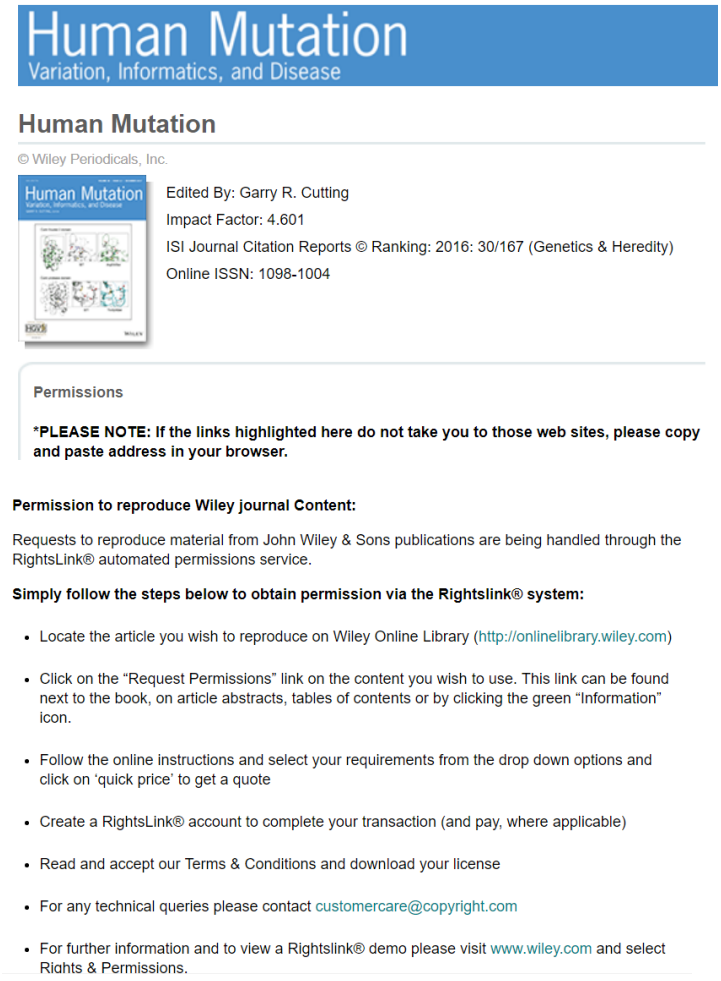
under normal culture conditions but are highly susceptible to free cholesterol-induced death. Molecular genetic evidence that the induction of phosphatidylcholine biosynthesis in free cholesterol-loaded macrophages is an adaptive response. *The Journal of Biological Chemistry*, *275*(45), 35368–76. https://doi.org/10.1074/jbc.M007099200

Zhang, Y., Luoh, S.-M., Hon, L. S., Baertsch, R., Wood, W. I., & Zhang, Z. (2007). GeneHub-GEPIS: digital expression profiling for normal and cancer tissues based on an integrated gene database. *Nucleic Acids Research*, *35*(Web Server issue), W152-8. https://doi.org/10.1093/nar/gkm381

Permissions

Chapter 2 is reprinted from the final accepted version of the journal article:

Jurgens J, Sobreira N, Modaff P, Reiser CA, Seo SH, Seong MW, Park SS, Kim OH, Cho TJ, Pauli RM. (2015). Novel *COL2A1* variant (c.619G>A, p.Gly207Arg) manifesting as a phenotype similar to progressive pseudorheumatoid dysplasia and spondyloepiphyseal dysplasia, Stanescu type. *Hum Mutat.* 36(10): 1004-8.



Human Mutation
Variation, Informatics, and Disease

Human Mutation

Edited By: Garry R. Cutting
Impact Factor: 4.601
ISI Journal Citation Reports © Ranking: 2016: 30/167 (Genetics & Heredity)
Online ISSN: 1098-1004

Permissions

***PLEASE NOTE: If the links highlighted here do not take you to those web sites, please copy and paste address in your browser.**

Permission to reproduce Wiley journal Content:

Requests to reproduce material from John Wiley & Sons publications are being handled through the RightsLink® automated permissions service.

Simply follow the steps below to obtain permission via the Rightslink® system:

- Locate the article you wish to reproduce on Wiley Online Library (http://onlinelibrary.wiley.com)
- Click on the "Request Permissions" link on the content you wish to use. This link can be found next to the book, on article abstracts, tables of contents or by clicking the green "Information" icon.
- Follow the online instructions and select your requirements from the drop down options and click on 'quick price' to get a quote
- Create a RightsLink® account to complete your transaction (and pay, where applicable)
- Read and accept our Terms & Conditions and download your license
- For any technical queries please contact customercare@copyright.com
- For further information and to view a Rightslink® demo please visit www.wiley.com and select Rights & Permissions.

AUTHORS - If you wish to reuse your own article (or an amended version of it) in a new publication of which you are the author, editor or co-editor, prior permission is not required (with the usual acknowledgements). However, a formal grant of license can be downloaded free of charge from RightsLink by selecting “Author of this Wiley article” as your requestor type.

Individual academic authors who are wishing to reuse up to 3 figures or up to 400 words from this journal to republish in a new journal article they are writing should select **University/Academic** as the requestor type. They will then be able to download a free permission license.

Either of the above who are publishing a new journal article or book chapter with an **STM Signatory Publisher** may also select that requestor type and the STM Signatory publisher’s name from the resulting drop-down list in RightsLink. This list is regularly updated. The requestor is required to complete the republication details, including the publisher name, during the request process. They will then be able to download a free permissions license.

Photocopying

Teaching institutions with a current paid subscription to the journal may make multiple copies for teaching purposes without charge, provided such copies are not resold or copied. In all other cases, permission should be obtained from a reproduction rights organisation (see below) or directly from RightsLink®.

Copyright Licensing Agency

Institutions based in the UK with a valid photocopying and/or digital license with the Copyright Licensing Agency may copy excerpts from Wiley books and journals under the terms of their license. For further information go to CLA.

Copyright Clearance Center

Institutions based in the US with a valid photocopying and/or digital license with the Copyright Clearance Center may copy excerpts from Wiley books and journals under the terms of their license, please go to CCC.

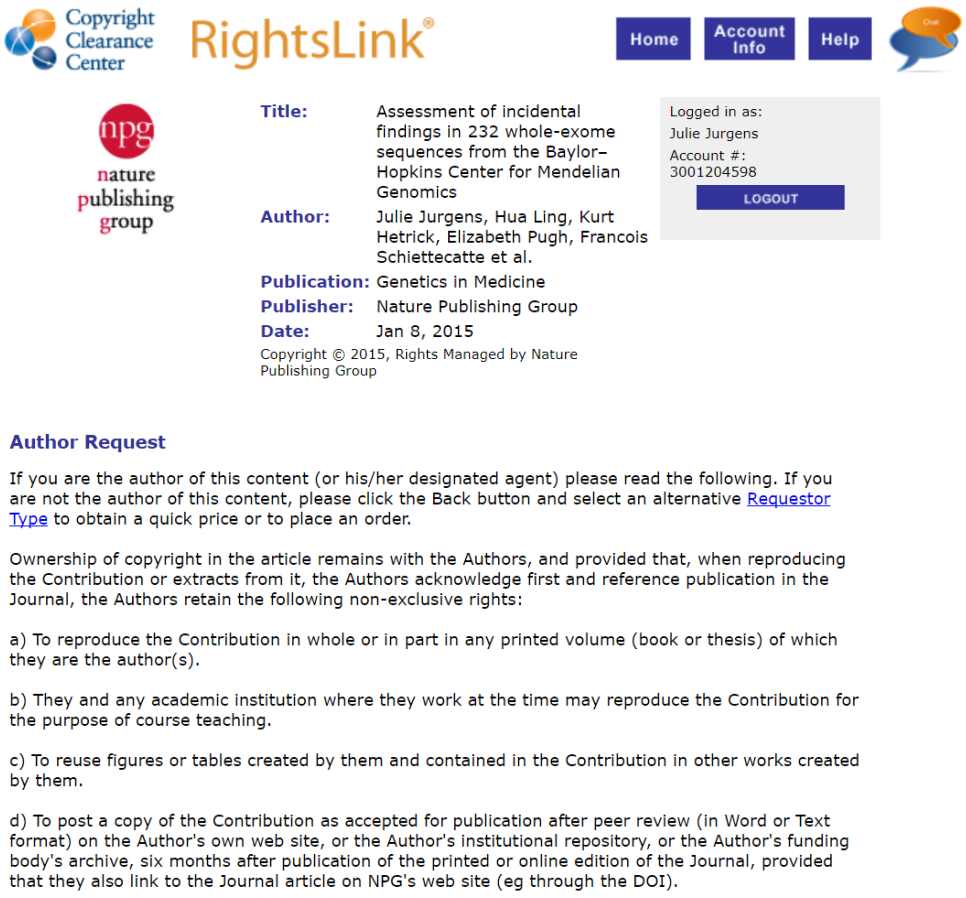
Other Territories: Please contact your local reproduction rights organisation.

For further information please visit www.wiley.com and select Rights & Permissions.

If you have any questions about the permitted uses of a specific article, please contact us at permissions@wiley.com

Chapter 3 is reprinted from the final accepted journal article:

Jurgens J, Ling H, Hetrick K, Pugh E, Schiettecatte F, Doheny K, Hamosh A, Avramopoulos D, Valle D, Sobreira N. (2015). Assessment of incidental findings in 232 whole exome sequences from the Baylor-Hopkins Center for Mendelian Genomics. *Genet Med*. 17(10): 782-8.



Copyright Clearance Center RightsLink®

Home | Account Info | Help

nature publishing group

Title: Assessment of incidental findings in 232 whole-exome sequences from the Baylor–Hopkins Center for Mendelian Genomics

Author: Julie Jurgens, Hua Ling, Kurt Hetrick, Elizabeth Pugh, Francois Schiettecatte et al.

Publication: Genetics in Medicine

Publisher: Nature Publishing Group

Date: Jan 8, 2015

Copyright © 2015, Rights Managed by Nature Publishing Group

Logged in as:
Julie Jurgens
Account #:
3001204598
LOGOUT

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select an alternative Requestor Type to obtain a quick price or to place an order.

Ownership of copyright in the article remains with the Authors, and provided that, when reproducing the Contribution or extracts from it, the Authors acknowledge first and reference publication in the Journal, the Authors retain the following non-exclusive rights:

a) To reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

b) They and any academic institution where they work at the time may reproduce the Contribution for the purpose of course teaching.

c) To reuse figures or tables created by them and contained in the Contribution in other works created by them.

d) To post a copy of the Contribution as accepted for publication after peer review (in Word or Text format) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the Journal article on NPG's web site (eg through the DOI).

NPG encourages the self-archiving of the accepted version of your manuscript in your funding agency's or institution's repository, six months after publication. This policy complements the recently announced policies of the US National Institutes of Health, Wellcome Trust and other research funding bodies around the world. NPG recognises the efforts of funding bodies to increase access to the research they fund, and we strongly encourage authors to participate in such efforts.

Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read NPG's online author reuse guidelines.

For full paper portion: Authors of original research papers published by NPG are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v2.0

BACK CLOSE WINDOW

Chapter 4 is reprinted from the final accepted version of the journal article:

Hoover-Fong J, Sobreira N, Jurgens J, Modaff P, Blout C, Moser A, Kim O-H, Cho T-J, Cho SY, Kim SJ, Jin D-K, Kitoh H, Park W-Y, Ling H, Hetrick KN, Doheny KF, Valle D, Pauli RM. (2014). Mutations in *PCYT1A*, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy. *Am J Hum Genet*. 94(1): 105-12.

Copyright Clearance Center RightsLink® Account Info Help

Title: Mutations in PCYT1A, Encoding a Key Regulator of Phosphatidylcholine Metabolism, Cause Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy

Author: Julie Hoover-Fong,Nara Sobreira,Julie Jurgens,Peggy Modaff,Carrie Blout,Ann Moser,Ok-Hwa Kim,Tae-Joon Cho,Sung Yoon Cho,Sang Jin Kim,Dong-Kyu Jin,Hiroshi Kitoh,Woong-Yang Park,Hua Ling,Kurt N. Hetrick,Kimberly F. Doheny,David Valle,Richard M. Pauli

Publication: The American Journal of Human Genetics

Publisher: Elsevier

Date: 2 January 2014

Copyright © 2014, Elsevier

Logged in as:
Julie Jurgens
Account #:
3001204598
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Julie A Jurgens ("You") and Elsevier ("Elsevier") consists of your order details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License number	Reference confirmation email for license number
License date	Dec, 12 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	The American Journal of Human Genetics
Licensed Content Title	Mutations in PCYT1A, Encoding a Key Regulator of Phosphatidylcholine Metabolism, Cause Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy
Licensed Content Author	Julie Hoover-Fong,Nara Sobreira,Julie Jurgens,Peggy Modaff,Carrie Blout,Ann Moser,Ok-Hwa Kim,Tae-Joon Cho,Sung Yoon Cho,Sang Jin Kim,Dong-Kyu Jin,Hiroshi Kitoh,Woong-Yang Park,Hua Ling,Kurt N. Hetrick,Kimberly F. Doheny,David Valle,Richard M. Pauli
Licensed Content Date	2 January 2014
Licensed Content Volume	94
Licensed Content Issue	1
Licensed Content Pages	8
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes

Will you be translating?	No
Title of your thesis/dissertation	Explaining rare Mendelian phenotypes: Exome sequencing and functional analysis of spondylometaphyseal dysplasia with cone-rod dystrophy
Expected completion date	Dec 2017
Estimated size (number of pages)	250
Requestor Location	Julie Jurgens 733 N. Broadway MRB 512 BALTIMORE, MD 21205 United States Attn: Julie Jurgens
Publisher Tax ID	98-0397604
Billing Type	Invoice
Billing address	Julie Jurgens 733 N. Broadway MRB 512 BALTIMORE, MD 21205 United States Attn: Julie Jurgens
Total	0.00 USD

CLOSE WINDOW

ELSEVIER LICENSE TERMS AND CONDITIONS

Dec 12, 2017

This Agreement between Julie Jurgens ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

The publisher has provided special terms related to this request that can be found at the end of the Publisher's Terms and Conditions.

License Number
4246521139253
License date
Dec 12, 2017
Licensed Content Publisher
Elsevier
Licensed Content Publication
The American Journal of Human Genetics
Licensed Content Title
Mutations in PCYT1A, Encoding a Key Regulator of Phosphatidylcholine Metabolism, Cause Spondylometaphyseal Dysplasia with Cone-Rod Dystrophy
Licensed Content Author
Julie Hoover-Fong,Nara Sobreira,Julie Jurgens,Peggy Modaff,Carrie Blout,Ann Moser,Ok-Hwa Kim,Tae-Joon Cho,Sung Yoon Cho,Sang Jin Kim,Dong-Kyu Jin,Hiroshi Kitoh,Woong-Yang Park,Hua Ling,Kurt N. Hetrick,Kimberly F. Doheny,David Valle,Richard M. Pauli
Licensed Content Date

Jan 2, 2014
Licensed Content Volume
94
Licensed Content Issue
1
Licensed Content Pages
8
Start Page
105
End Page
112
Type of Use
reuse in a thesis/dissertation
Portion
full article
Format
both print and electronic
Are you the author of this Elsevier article?
Yes
Will you be translating?
No
Title of your thesis/dissertation
Explaining rare Mendelian phenotypes: Exome sequencing and functional analysis of spondylometaphyseal dysplasia with cone-rod dystrophy
Expected completion date
Dec 2017
Estimated size (number of pages)
250
Requestor Location
Julie Jurgens
733 N. Broadway
MRB 512

BALTIMORE, MD 21205
United States
Attn: Julie Jurgens
Publisher Tax ID
98-0397604
Billing Type
Invoice
Billing Address
Julie Jurgens
733 N. Broadway
MRB 512

BALTIMORE, MD 21205

United States
Attn: Julie Jurgens
Total
0.00 USD
Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never

granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include

permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.
Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.
17. **For journal authors:** the following clauses are applicable in addition to the above:
Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).
Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).
If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.
Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.
Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.
Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.
19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing.

20. **Other Conditions**: Permission is granted to submit your article in print and electronic format. This license permits you to post this Elsevier article online on your Institution's website if the content is embedded within your thesis.

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Julie A. Jurgens
Curriculum Vitae

1518 Park Ave, Apt 210N
Baltimore, MD 21217
217.433.9589
jurgens88@gmail.com

707 N. Broadway, Miller Research Building 512
Baltimore, MD 21205
410.614.2577
jjurgen2@jhmi.edu

EDUCATION

Johns Hopkins University School of Medicine August 2012-Dec 2017
Baltimore, MD
Ph.D. Candidate, Predoctoral Training Program in Human Genetics

University of Illinois at Urbana-Champaign August 2006-May 2010
Champaign, IL
Bachelor of Science in Molecular and Cell Biology, Chemistry Minor
GPA: 3.62/4.00

PROFESSIONAL EXPERIENCE

Johns Hopkins School of Medicine Baltimore, MD Aug 2012-Dec 2017
Predoctoral Training Program in Human Genetics
Visual Neuroscience Training Program
Lab of David Valle, M.D.
PhD Candidate

- Developed and characterized model systems to elucidate the pathophysiological mechanisms underlying spondylometaphyseal dysplasia with cone-rod dystrophy.
- Investigated novel genetic bases of unexplained Mendelian disorders through next-generation sequencing analysis at the Baylor-Hopkins Center for Mendelian Genomics.
- Secured independent research funding to conduct a trainee-directed whole genome sequencing project in collaboration with a fellow graduate student. Won an independent fellowship to obtain financial support and pursue educational opportunities in visual neuroscience.
- Delivered written and oral presentations, including manuscripts for peer review.
- Supervised and mentored undergraduate trainees and junior graduate students.

Children's National Medical Center Washington, DC June 2010 – Feb 2012
University of Maryland School of Medicine Baltimore, MD Feb 2012-Aug 2012
Lab of Pedro Jose, Ph.D.
Research Technician

- Studied mechanisms underlying essential hypertension through a number of collaborative projects. Supervised breeding, maintenance, and genotyping of a large mouse colony and helped coordinate shipment to a new institution.
- Presented results on a weekly basis, co-authored publications, and mentored three high school students.

University of Illinois at Urbana-Champaign Champaign, IL May 2008 – May 2010
Lab of David Clayton, Ph.D.

Undergraduate Research Assistant

- Investigated modulations in gene expression associated with auditory learning in zebra finches. Completed a senior thesis which was awarded High Distinction and communicated findings verbally and orally.

University of Illinois at Urbana-Champaign Champaign, IL Sept 2009 – May 2010
Lab of Rebecca Krisher, Ph.D.

Undergraduate Laboratory Assistant

- Studied metabolic and reproductive features of Ossabaw pigs in relation to Polycystic Ovarian Syndrome (PCOS).
- Worked with a team to perform ultrasounds, obtain body measurements, collect blood and tissue samples, and care for animals.

PUBLICATIONS

Florwick A, Dharmaraj T, **Jurgens J**, Valle D, Wilson KL. (2017). *LMNA* sequences of 60,706 individuals link the p.G602S variant to Type 2 Diabetes and reveal novel missense variants shared by lamin A and C, or unique to Prelamin A. *Front Genet*. 8: 79.

Schossig A, Bloch-Zupan A, Lussi A, Wolf NI, Raskin S, Cohen M, Giuliano F, **Jurgens J**, Krabichler B, Koolen DA, Sobreira NL, Maurer E, Muller-Bolla M, Penzien J, Zschocke J, Kapferer-Seebacher I. (2017). *SLC13A5* is the second gene associated with Kohlschütter-Tönz syndrome. *J Med Genet*. 54(1):54-62.

You J, Sobreira NL, Grange DK, Gable D, **Jurgens J**, Belnap N, Siniard A, Szelinger S, Schrauwen I, Richholt RF, Vallee SE, Palko Dinulos MB, Valle D, Armanios M, Hoover-Fong J. (2016). A Newly Recognized Intellectual Disability Disorder Caused by Variants in *TELO2,* a gene encoding a component of the TTT complex. *Am J Hum Genet*. 98(5): 909-918.

Jurgens J, Sobreira N, Modaff P, Reiser CA, Seo SH, Seong MW, Park SS, Kim OH, Cho TJ, Pauli RM. (2015). Novel *COL2A1* variant (c.619G>A, p.Gly207Arg) manifesting as a phenotype similar to progressive pseudorheumatoid dysplasia and spondyloepiphyseal dysplasia, Stanescu type. *Hum Mutat.* 36(10): 1004-8.

Jurgens J, Ling H, Hetrick K, Pugh E, Schiettecatte F, Doheny K, Hamosh A, Avramopoulos D, Valle D, Sobreira N. (2015). Assessment of incidental findings in 232 whole exome sequences from the Baylor-Hopkins Center for Mendelian Genomics. *Genet Med*. 17(10): 782-8.

Minillo RM, Sobreira N, de Faria Soares Mde F, **Jurgens J**, Ling H, Hetrick KN, Doheny KF, Valle D, Brunoni D, Perez AB. (2014). Novel Deletion of SERPINF1 Causes Autosomal Recessive Osteogenesis Imperfecta Type VI in Two Brazilian Families. *Mol Syndromol*. 5(6):268-75.

Ennis RC, Asico LD, Armando I, Yang J, Feranil JB, **Jurgens JA**, Escano CS Jr, Yu P, Wang X, Sibley DR, Jose PA, Villar VA. (2014). Dopamine D1-like receptors regulate the α1A adrenergic receptor in human renal proximal tubule cells and D1-like dopamine receptor knockout mice. *Am J Physiol Renal Physiol.* 307(11): F1238-48.

Hoover-Fong J, Sobreira N, **Jurgens J**, Modaff P, Blout C, Moser A, Kim O-H, Cho T-J, Cho SY, Kim SJ, Jin D-K, Kitoh H, Park W-Y, Ling H, Hetrick KN, Doheny KF, Valle D, Pauli RM. (2014). Mutations in *PCYT1A*, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy. *Am J Hum Genet*. 94(1): 105-12.

Villar VA, Armando I, Sanada H, Frazer LC, Russo CM, Notario PM, Lee H, Comisky L, Russell HA, Yang Y, **Jurgens JA**, Jose PA, Jones JE. (2013). Novel role of sorting nexin 5 in renal D(1) dopamine receptor trafficking and function: implications for hypertension. *FASEB J*. 27(5):1808-19.

Jones JE, **Jurgens JA**, Evans SA, Ennis RC, Villar VA, Jose PA. (2012). Mechanisms of fetal programming in hypertension. *Int J Pediatr*. 2012:584831.

LEADERSHIP AND VOLUNTEER EXPERIENCE

American Society of Human Genetics (ASHG) Bethesda, MD Jan 2015-Dec 2017
Training and Development Committee

- Work with a team of ASHG trainees and staff to develop events and resources for trainees.
- Conceptualized and generated a quarterly newsletter for ASHG trainees. Lead a working group in which I contribute and edit original articles, manage a group of trainee writers, and finalize content with ASHG staff.
- Developed an initiative to make informational videos about ASHG resources. Led working groups in 2015 and 2016 to script, narrate, film, and edit videos.
- Plan career development events for annual ASHG meetings. Invite speakers, coordinate session logistics, and moderate career development sessions.
- Review annual trainee applications for all ASHG committees.

Predoctoral Training Program in Human Genetics Jan 2015-Jan 2017

Johns Hopkins University
Student Representative

- Plan and execute annual student orientation and recruitment events.
- Chaired a committee of students for selection and invitation of a speaker for the 2015 Barton Childs Memorial Lectureship. Coordinated lecture logistics and introduced the speaker.
- Helped to organize and co-moderate the 2017 Human Genetics Career Day.
- Collaborate with program administrators to develop new program initiatives, deliver information to the student body, and collect feedback to improve departmental resources and events.

Online Mendelian Inheritance in Man (OMIM) October 2015-June 2016
Johns Hopkins University
Volunteer Article Writer

- Contributed regular articles to OMIM describing genes and their relation to human phenotypes.

American Society of Human Genetics, 2016 Annual Meeting October 2016
Co-moderator, "Diseases of the Nervous System" Platform Session

- With a fellow trainee, managed session timing, speaker introductions, and question/answer period. Coordinated logistics with presenters and ASHG's Program Committee prior to the meeting.

Leadership Alliance National Symposium 2016 July 2016
Session Moderator and Poster Presentation Judge

- Reviewed and provided feedback on abstracts, posters, and oral presentations for undergraduate symposium attendees.
- Moderated a session for undergraduate presenters. Coordinated logistics prior to the meeting and managed introductions, timing, and question/answer period.

Molecular and Cell Biology Advising University of Illinois Oct 2008–May 2010
Undergraduate Research Information Workshop Leader

- Advised 10-20 undergraduates per month on how to become involved in campus research.
- Led group discussions, delivered presentations, and collaborated with others to develop departmental events.

Future Medical Research Investigators University of Illinois Aug 2008 – May 2010
Treasurer, August 2009-May 2010

- Informed students of the dual MD/PhD degree option, how to pursue it, and how to use available resources to learn more about programs.
- Invited university affiliates and external guest speakers to deliver presentations.
- Managed accounts and delivered presentations to the student funding board to raise finances.

Illini Medical Screening Society University of Illinois Jan 2009 – May 2010
General Member

- Provided free health education and screenings to local low-income or uninsured individuals.
- Measured blood pressure and metabolites to determine participants' likelihood of developing hypertension, diabetes, and hyperlipidemia.
- Helped to organize screenings and recruit local doctors to interpret test results and guide participants toward additional resources.

AWARDS AND HONORS

Schepens Eye Research Institute, Massachusetts Eye and Ear Hospital, Harvard Medical School

2018 Molecular Bases of Eye Diseases Postdoctoral Fellowship Recipient

Burroughs-Wellcome Fund/ Maryland Genetics, Epidemiology, and Medicine Program

2016 Wolfe Street Competition Award Co-Recipient for proposal to detect novel genetic mechanisms of strabismus through whole genome sequencing
The Wolfe Street Competition provides $15,000 for an independent research project led by two trainees from the Johns Hopkins Bloomberg School of Public Health and the Johns Hopkins University School of Medicine to foster collaboration between laboratory and population scientists.

Johns Hopkins University Wilmer Eye Institute and Department of Neuroscience

2015 Visual Neuroscience Training Program Graduate Fellowship Recipient

American Society of Human Genetics

2013 ASHG/ Charles J. Epstein Trainee Award for Excellence in Human Genetics Research, Semifinalist

University of Illinois at Urbana-Champaign

2010 Cell & Developmental Biology Department Award for Distinguished Research-Best Senior Thesis
2010 Graduation with High Distinction

TEACHING AND MENTORSHIP

Graduate Course in Advanced Topics in Human Genetics Nov 2014-May 2015
Johns Hopkins University
Teaching Assistant

- Worked with course instructors and a fellow graduate student to shape course content, manage the course, and answer students' questions.
- Led weekly discussion sections, wrote and graded assignments and exams, developed review materials, and coordinated course logistics.

Predoctoral Training Program in Human Genetics Sept 2015-Dec 2017
Johns Hopkins University

Graduate and undergraduate student mentor

- Mentor undergraduate and junior graduate students in their laboratory rotations and thesis projects.
- Develop projects for students and supervise progress in lab.
- Help students design, conduct, and troubleshoot experiments. Assist them in developing abstracts, posters, and oral presentations. Provide ongoing resources for scientific and career development.

Thomas Jefferson High School for Science and Technology June 2010-Aug 2012
McLean, Virginia
High school student mentor

- Taught high school students techniques in molecular biology, supervised laboratory progress, and offered troubleshooting assistance.
- Helped students develop written and oral presentations, write scientifically, edit college applications, and navigate educational and career development.

Mentees:

High school students: Alan Barte, Rishav Adhikari, Riley Ennis
Undergraduate students: Adam Uppendahl, Sophie Nguyen
Graduate students: Arianna Franca, Sarah Robbins, Anh-Thu Lam

CONFERENCE PRESENTATIONS

Assessment of incidental findings in whole exome sequencing data from the Baylor-Hopkins Center for Mendelian Genomics. Platform Presentation. American Society of Human Genetics 2013.

Abnormal phospholipid metabolism due to variants in *PCYT1A* causes spondylometaphyseal dysplasia with cone-rod dystrophy. Poster Presentation. Hopkins MD-GEM Genetics Research Day 2014.

Assessment of incidental findings in whole exome sequencing data from the Baylor-Hopkins Center for Mendelian Genomics. Poster Presentation. Hopkins MD-GEM Genetics Research Day 2014.

Abnormal phospholipid metabolism due to variants in *PCYT1A* causes spondylometaphyseal dysplasia with cone-rod dystrophy. Poster Presentation. American Society of Human Genetics 2014.

Strategy for Generating and Characterizing a Zebrafish Knockout Model of Spondylometaphyseal Dysplasia with Cone-rod Dystrophy. Poster Presentation. Hopkins MD-GEM Genetics Research Day 2015.

Strategy for Generating and Characterizing a Zebrafish Knockout Model of Spondylometaphyseal Dysplasia with Cone-rod Dystrophy. Poster Presentation. American Society of Human Genetics 2015.

Phenotypic characterization of cell lines from patients with spondylometaphyseal dysplasia with cone-rod dystrophy (SMD-CRD). Poster Presentation. American Society of Human Genetics 2016.

Investigation of novel genetic bases for strabismus using linkage analysis and next-generation sequencing. Hopkins MD-GEM Genetics Research Day 2017.

Investigation of novel genetic bases for strabismus using linkage analysis and next-generation sequencing. Poster Presentation. American Society of Human Genetics 2017.

PROFESSIONAL SOCIETY MEMBERSHIPS

Member, American Society of Human Genetics June 2013-Present